612,314 CITED ON 1444

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12N 15/12, A61K 37/02 C07K 13/00, C12N 1/21 // (C12N 1/21, C12R 1:19)

A1

(11) International Publication Number: WO 94/00571

(43) International Publication Date: 6 January 1994 (06.01.94)

(21) International Application Number: PCT/GB93/01282

(22) International Filing Date: 16 June 1993 (16.06.93)

(30) Priority data:

9213376.8	24 June 1992 (24.06.92)	GB
9304057.4	1 March 1993 (01.03.93)	GB

(71) Applicant (*for all designated States except US*): SMITHKLINE BEECHAM PLC [GB/GB]; New Horizons Court, Brentford, Middlesex TW8 9EP (GB).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*) : SMITH, Richard, Anthony, Godwin [GB/GB]; DODD, Ian [GB/GB]; FREEMAN, Anne, Mary [GB/GB]; MOSSAKOWSKA, Danuta, Eva, Irena [GB/GB]; SmithKline Beecham Pharmaceuticals, Great Burgh, Yew Tree Bottom Road, Epsom, Surrey KT18 5XQ (GB).

(74) Agent: VALENTINE, Jill, Barbara; SmithKline Beecham, Corporate Patents, Great Burgh, Yew Tree Bottom Road, Epsom, Surrey KT18 5XQ (GB).

(81) Designated States: JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: SOLUBLE CR1 DERIVATIVES

(57) Abstract

A soluble polypeptide comprising, in sequence, one to four short consensus repeats (SCR) selected from SCR 1,2,3 and 4 of long homologous repeat A(LHR-A) as the only structurally and functionally intact SCR domains of CR1 and including at least SCR3.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NE	Niger
BE	Belgium	GN	Guinea	NL	Netherlands
BF	Burkina Faso	GR	Greece	NO	Norway
BG	Bulgaria	HU	Hungary	NZ	New Zealand
BJ	Benin	IE	Ireland	PL	Poland
BR	Brazil	IT	Italy	PT	Portugal
BY	Belarus	JP	Japan	RO	Romania
CA	Canada	KP	Democratic People's Republic of Korea	RU	Russian Federation
CF	Central African Republic			SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovak Republic
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	UA	Ukraine
DE	Germany	MG	Madagascar	US	United States of America
DK	Denmark	ML	Mali	UZ	Uzbekistan
ES	Spain	MN	Mongolia	VN	Viet Nam
FI	Finland				

SOLUBLE CR1 DERIVATIVES.

The present invention relates to polypeptides and their use in the diagnosis and therapy of disorders involving complement activity and various inflammatory and immune disorders.

Constituting about 10% of the globulins in normal serum, the complement system is composed of many different proteins that are important in the immune system's response to foreign antigens. The complement system becomes activated when its primary components are cleaved and the products alone or with other proteins, activate additional complement proteins resulting in a proteolytic cascade. Activation of the complement system leads to a variety of responses including increased vascular permeability, chemotaxis of phagocytic cells, activation of inflammatory cells, opsonization of foreign particles, direct killing of cells and tissue damage. Activation of the complement system may be triggered by antigen-antibody complexes (the classical pathway) or, for example, by lipopolysaccharides present in cell walls of pathogenic bacteria (the alternative pathway).

Complement receptor type 1 (CR1) has been shown to be present on the membranes of erythrocytes, monocytes/macrophages, granulocytes, B cells, some T cells, splenic follicular dendritic cells, and glomerular podocytes. CR1 binds to the complement components C3b and C4b and has also been referred to as the C3b/C4b receptor. The structural organisation and primary sequence of one allotype of CR1 is known (Klickstein *et al.*, 1987, J. Exp. Med. 165:1095-1112, Klickstein *et al.*, 1988, J. Exp. Med. 168:1699-1717; Hourcade *et al.*,1988, J. Exp. Med. 168:1255-1270, WO 89/09220, WO 91/05047). It is composed of 30 short consensus repeats (SCRs) that each contain around 60-70 amino acids. In each SCR, around 29 of the average 65 amino acids are conserved. Each SCR has been proposed to form a three dimensional triple loop structure through disulphide linkages with the third and first and the fourth and second half-cystines in disulphide bonds. CR1 is further arranged as 4 long homologous repeats (LHRs) of 7 SCRs each. Following a leader sequence, the CR1 molecule consists of the N-terminal LHR-A, the next two repeats, LHR-B and LHR-C, and the most C-terminal LHR-D followed by 2 additional SCRs, a 25 residue putative transmembrane region and a 43 residue cytoplasmic tail.

Based on the mature CR1 molecule having a predicted N-terminal glutamine residue, hereinafter designated as residue 1, the first four SCR domains of LHR-A are defined herein as consisting of residues 2-58, 63-120, 125-191 and 197-252, respectively, of mature CR1.

Hourcade *et al.*,1988, J. Exp. Med. 168:1255-1270 observed an alternative polyadenylation site in the human CR1 transcriptional unit that was predicted to

produce a secreted form of CR1. The mRNA encoded by this truncated sequence comprises the first 8.5 SCRs of CR1, and encodes a protein of about 80 kDa which was proposed to include the C4b binding domain. When a cDNA corresponding to this truncated sequence was transfected into COS cells and expressed, it demonstrated the expected C4b binding activity but did not bind to C3b (Krych *et al.*,1989, FASEB J. 3:A368; Krych *et al.* Proc. Nat. Acad. Sci. 1991, 88, 4353-7). Krych *et al.*, also observed a mRNA similar to the predicted one in several human cell lines and postulated that such a truncated soluble form of CR1 with C4b binding activity may be synthesised in humans.

In addition, Makrides *et al.* (1992, J. Biol. Chem. 267 (34) 24754-61) have expressed SCR 1 + 2 and 1 + 2 + 3 + 4 of LHR-A as membrane-attached proteins in CHO cells.

Several soluble fragments of CR1 have also been generated via recombinant DNA procedures by eliminating the transmembrane region from the DNAs being expressed (WO 89/09220, WO 91/05047). The soluble CR1 fragments were functionally active, bound C3b and/or C4b and demonstrated Factor I cofactor activity depending upon the regions they contained. Such constructs inhibited *in vitro* complement-related functions such as neutrophil oxidative burst, complement mediated hemolysis, and C3a and C5a production. A particular soluble construct, sCR1/pBSCR1c, also demonstrated *in vivo* activity in a reversed passive Arthus reaction (WO 89/09220, WO 91/05047; Yeh *et al.*, 1991, J. Immunol. 146:250), suppressed post-ischemic myocardial inflammation and necrosis (WO 89/09220, WO 91/05047; Weisman *et al.*, Science, 1990, 249:146-1511; Dupe, R. *et al.* Thrombosis & Haemostasis (1991) 65(6) 695.) and extended survival rates following transplantation (Pruitt & Bollinger, 1991, J. Surg. Res 50:350; Pruitt *et al.*, 1991 Transplantation 52; 868). Furthermore, co-formulation of sCR1/pBSCR1c with p-anisoylated human plasminogen-streptokinase-activator complex (APSAC) resulted in similar anti-haemolytic activity as sCR1 alone, indicating that the combination of the complement inhibitor sCR1 with a thrombolytic agent was feasible (WO 91/05047).

Soluble polypeptides corresponding to part of CR1 have now been found to possess functional complement inhibitory, including anti-haemolytic, activity.

According to the present invention there is provided a soluble polypeptide comprising, in sequence, one to four short consensus repeats (SCR) selected from SCR 1, 2, 3 and 4 of long homologous repeat A (LHR-A) as the only structurally and functionally intact SCR domains of CR1 and including at least SCR3.

In preferred aspects, the polypeptide comprises, in sequence, SCR 1, 2, 3 and 4 of LHR-A or SCR 1, 2 and 3 of LHR-A as the only structurally and functionally

intact SCR domains of CR1.

It is to be understood that variations in the amino acid sequence of the polypeptide of the invention by way of addition, deletion or conservative substitution of residues, including allelic variations, in which the biological activity of the polypeptide is retained, are encompassed by the invention. Conservative substitution is understood to mean the retention of the charge and size characteristics of the amino acid side chain, for example arginine replaced by histidine.

In one aspect, the polypeptide of the invention may be represented symbolically as follows:

$$NH_2\text{-}V^1\text{-SCR1-}W^1\text{-SCR2-}X^1\text{-SCR3-}Y^1\text{-OH} \qquad (I)$$

in which SCR1 represents residues 2-58 of mature CR1, SCR2 represents residues 63-120 of mature CR1, SCR3 represents residues 125-191 of mature CR1, and V^1, W^1, X^1 and Y^1 represent bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

In a preferred embodiment of formula (I), W^1, X^1 and Y^1 represent residues 59-62, 121-124 and 192-196, respectively, of mature CR1 and V^1 represents residue 1 of mature CR1 optionally linked via its N-terminus to methionine.

In another aspect the polypeptide of the invention may be represented symbolically as follows:

$$NH_2\text{-}V^2\text{-SCR1-}W^2\text{-SCR2-}X^2\text{-SCR3-}Y^2\text{-SCR4-}Z^2\text{OH} \qquad (II)$$

in which SCR1, SCR2 and SCR3 are as hereinbefore defined, SCR4 represents residues 197-252 of mature CR1 and V^2, W^2, X^2, Y^2 and Z^2 represents bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

In preferred embodiments of formula (II), W^2, X^2, Y^2 and Z^2 represent residues 59-62, 121-124, 192-196, and residues 253 respectively, of mature CR1 and V^2 represents residue 1 of mature CR1 optionally linked via its N-terminus to methionine.

In one particular embodiment of formula (II) arginine 235 is replaced by histidine.

In the preferred embodiment of formula (II), residue 235 is arginine.

In one further aspect, the polypeptide of the invention may be represented symbolically as follows:

$$NH_2\text{-}X^3\text{-SCR3-}Y^3\text{-OH} \qquad (III)$$

in which SCR3 is as hereinbefore defined and X^3 and Y^3 represent bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

In a preferred embodiment of formula (III) X^3 represents amino acids 122-124 of mature CR1 optionally linked to methionine at its N-terminus and Y^4 represents amino acids 192-196 of mature CR1.

In another further aspect, the polypeptide of the invention may be represented symbolically as follows:

$$NH_2\text{-}X^4\text{-SCR3-}Y^4\text{-SCR4-}Z^4\text{-OH} \qquad \text{(IV)}$$

in which SCR3 and SCR4 are as hereinbefore defined and X^4, Y^4 and Z^4 represent bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

In a preferred embodiment of formula (IV) X^4 represents amino acids 122-124 of mature CR1 optionally linked to methionine at its N-terminus and Y^4 and Z^4 represent amino acids 192-196 and 253 respectively of mature CR1.

In a further aspect, the invention provides a process for preparing a CR1 polypeptide according to the invention which process comprises expressing DNA encoding said polypeptide in a recombinant host cell and recovering the product.

In particular, the process may comprise the steps of:

i) preparing a replicable expression vector capable, in a host cell, of expressing a DNA polymer comprising a nucleotide sequence that encodes said polypeptide;

ii) transforming a host cell with said vector;

iii) culturing said transformed host cell under conditions permitting expression of said DNA polymer to produce said polypeptide; and

iv) recovering said polypeptide.

The DNA polymer comprising a nucleotide sequence that encodes the polypeptide also forms part of the invention.

The process of the invention may be performed by conventional recombinant techniques such as described in Sambrook *et al.*, Molecular Cloning : A laboratory manual 2nd Edition. Cold Spring Harbor Laboratory Press (1989) and DNA Cloning vols I, II and III (D. M. Glover ed.. IRL Press Ltd).

The invention also provides a process for preparing the DNA polymer by the condensation of appropriate mono-, di- or oligomeric nucleotide units.

The preparation may be carried out chemically, enzymatically, or by a combination of the two methods, *in vitro* or *in vivo* as appropriate. Thus, the DNA polymer may be prepared by the enzymatic ligation of appropriate DNA fragments, by conventional methods such as those described by D. M. Roberts *et al.*, in Biochemistry 1985, 24, 5090-5098.

The DNA fragments may be obtained by digestion f DNA containing the required sequences of nucleotides with appropriate restriction enzymes, by chemical synthesis, by enzymatic polymerisation, or by a combination f these methods.

Digestion with restriction enzymes may be performed in an appropriate buffer at a temperature of 20°-70°C, generally in a volume of 50µl or less with 0.1-10µg DNA.

Enzymatic polymerisation of DNA may be carried out *in vitro* using a DNA polymerase such as DNA polymerase 1 (Klenow fragment) in an appropriate buffer containing the nucleoside triphosphates dATP, dCTP, dGTP and dTTP as required at a temperature of 10°-37°C, generally in a volume of 50µl or less.

Enzymatic ligation of DNA fragments may be carried out using a DNA ligase such as T4 DNA ligase in an appropriate buffer at a temperature of 4°C to 37°C, generally in a volume of 50µl or less.

The chemical synthesis of the DNA polymer or fragments may be carried out by conventional phosphotriester, phosphite or phosphoramidite chemistry, using solid phase techniques such as those described in 'Chemical and Enzymatic Synthesis of Gene Fragments - A Laboratory Manual' (ed. H.G. Gassen and A. Lang), Verlag Chemie, Weinheim (1982), or in other scientific publications, for example M.J.Gait, H.W.D. Matthes M. Singh, B.S. Sproat and R.C. Titmas, Nucleic Acids Research, 1982, 10, 6243; B.S. Sproat and W. Bannwarth, Tetrahedron Letters, 1983, 24, 5771; M.D. Matteucci and M.H. Caruthers, Tetrahedron Letters, 1980, 21, 719; M.D. Matteucci and M.H. Caruthers, Journal of the American Chemical Society, 1981, 103, 3185; S.P. Adams *et al.*, Journal of the American Chemical Society, 1983, 105, 661; N.D. Sinha, J. Biernat, J. McMannus and H. Koester, Nucleic Acids Research, 1984, 12, 4539; and H.W.D. Matthes *et al.*, EMBO Journal, 1984, 3, 801. Preferably an automated DNA synthesiser (for example, Applied Biosystems 381A Synthesiser) is employed.

The DNA polymer is preferably prepared by ligating two or more DNA molecules which together comprise a DNA sequence encoding the polypeptide.

The DNA molecules may be obtained by the digestion with suitable restriction enzymes of vectors carrying the required coding sequences.

The precise structure of the DNA molecules and the way in which they are obtained depends upon the structure of the desired product. The design of a suitable strategy for the construction of the DNA molecule coding for the polypeptide is a routine matter for the skilled worker in the art.

In particular, consideration may be given to the codon usage of the particular host cell. The codons may be optimised for high level expression in *E. coli* using the principles set ut in Devereux *et al.*, (1984) Nucl. Acid Res., 12, 387.

The expression of the DNA polymer encoding the polypeptide in a recombinant host cell may be carried ut by means of a replicable xpression vector capable, in the host cell, of expressing the DNA polymer. The expression vector is

novel and also forms part of the invention.

The replicable expression vector may be prepared in accordance with the invention, by cleaving a vector compatible with the host cell to provide a linear DNA segment having an intact replicon, and combining said linear segment with one or more DNA molecules which, together with said linear segment, encode the polypeptide, under ligating conditions.

The ligation of the linear segment and more than one DNA molecule may be carried out simultaneously or sequentially as desired.

Thus, the DNA polymer may be preformed or formed during the construction of the vector, as desired. The choice of vector will be determined in part by the host cell, which may be prokaryotic, such as *E. coli*, or eukaryotic, such as mouse C127, mouse myeloma, chinese hamster ovary, fungi e.g. filamentous fungi or unicellular 'yeast' or an insect cell such as Drosophila. The host cell may also be in a transgenic animal. Suitable vectors include plasmids, bacteriophages, cosmids and recombinant viruses derived from, for example, baculoviruses or vaccinia.

The DNA polymer may be assembled into vectors designed for isolation of stable transformed mammalian cell lines expressing the fragment e.g. bovine papillomavirus vectors in mouse C127 cells, or amplified vectors in chinese hamster ovary cells (DNA Cloning Vol. II D.M. Glover ed. IRL Press 1985; Kaufman, R.J. *et al.*. Molecular and Cellular Biology 5, 1750-1759, 1985; Pavlakis G.N. and Hamer, D.H. Proceedings of the National Academy of Sciences (USA) 80, 397-401, 1983; Goeddel, D.V. *et al.*,European Patent Application No. 0093619, 1983).

The preparation of the replicable expression vector may be carried out conventionally with appropriate enzymes for restriction, polymerisation and ligation of the DNA, by procedures described in, for example, Sambrook *et al.*,cited above. Polymerisation and ligation may be performed as described above for the preparation of the DNA polymer. Digestion with restriction enzymes may be performed in an appropriate buffer at a temperature of 20°-70°C, generally in a volume of 50µl or less with 0.1-10µg DNA.

The recombinant host cell is prepared, in accordance with the invention, by transforming a host cell with a replicable expression vector of the invention under transforming conditions. Suitable transforming conditions are conventional and are described in, for example, Sambrook *et al.*, cited above, or "DNA Cloning" Vol. II, D.M. Glover ed., IRL Press Ltd, 1985.

The choice f transforming conditions is determined by the host cell. Thus, a bacterial host such as *E.coli*, may be treated with a solution of $CaCl_2$ (Cohen *et al.*,Proc. Nat. Acad. Sci., 1973, <u>69</u>, 2110) or with a soluti n comprising a mixture of RbCl, $MnCl_2$, potassium acetate and glycerol, and then with 3-[N-morpholino]-

propane-sulphonic acid, RbCl and glycerol or by electroporation as for example described by Bio-Rad Laboratories, Richmond, California, USA, manufacturers of an electroporator. Mammalian cells in culture may be transformed by calcium co-precipitation of the vector DNA onto the cells or by using cationic liposomes.

The invention also extends to a host cell transformed with a replicable expression vector of the invention.

Culturing the transformed host cell under conditions permitting expression of the DNA polymer is carried out conventionally, as described in, for example, Sambrook *et al.*, and "DNA Cloning" cited above. Thus, preferably the cell is supplied with nutrient and cultured at a temperature below 45°C.

The protein product is recovered by conventional methods according to the host cell. Thus, where the host cell is bacterial such as *E. coli* and the protein is expressed intracellularly, it may be lysed physically, chemically or enzymatically and the protein product isolated from the resulting lysate. Where the host cell is mammalian, the product is usually isolated from the nutrient medium.

Where the host cell is bacterial, such as *E. coli*, the product obtained from the culture may require folding for optimum functional activity. This is most likely if the protein is expressed as inclusion bodies. There are a number of aspects of the isolation and folding process that are regarded as important. In particular, the polypeptide is preferably partially purified before folding, in order to minimise formation of aggregates with contaminating proteins and minimise misfolding of the polypeptide. Thus, the removal of contaminating *E. coli* proteins by specifically isolating the inclusion bodies and the subsequent additional purification prior to folding are important aspects of the procedure.

The folding process is carried out in such a way as to minimise formation of intermediate-folded states of the polypeptide. Thus, careful consideration needs to be given to, among others, the salt type and concentration, temperature, protein concentration, redox buffer concentrations and duration of folding. The exact condition for any given polypeptide generally cannot be predicted and must be determined by experiment.

There are numerous methods available for the folding of proteins from inclusion bodies and these are known to the skilled worker in this field. The methods generally involve breaking all the disulphide bonds in the inclusion body, for example with 50mM 2-mercaptoethanol, in the presence of a high concentration of denaturant such as 8M urea or 6M guanidine hydrochloride. The next step is to remove these agents to allow folding of the proteins to occur. Formation of the disulphide bridges requires an oxidising environment and this may be provided in a number of ways, for example by air, or by incorporating a suitable redox system, for example a mixture of

reduced and oxidised glutathione.

Preferably, the inclusion body is solubilised using 8M urea, in the presence of mercaptoethanol, and protein is folded, after initial removal of contaminating proteins, by addition of cold buffer. A preferred buffer is 20mM ethanolamine containing 1mM reduced glutathione and 0.5mM oxidised glutathione. The folding is preferably carried out at a temperature in the range 1 to 5^{o}C over a period of 1 to 4 days.

If any precipitation or aggregation is observed, the aggregated protein can be removed in a number of ways, for example by centrifugation or by treatment with precipitants such as ammonium sulphate. Where either of these procedures are adopted, monomeric polypeptide is the major soluble product.

If the bacterial cell secretes the protein, folding is not usually necessary.

The polypeptide of this invention is useful in the treatment or diagnosis of many complement-mediated or complement-related diseases and disorders including, but not limited to, those listed below.

Disease and Disorders Involving Complement

Neurological Disorders

multiple sclerosis
stroke
Guillain Barré Syndrome
traumatic brain injury
Parkinson's disease
allergic encephalitis
Alzheimer's disease

Disorders of Inappropriate or Undesirable Complement Activation

haemodialysis complications
hyperacute allograft rejection
xenograft rejection
corneal graft rejection
interleukin-2 induced toxicity during IL-2 therapy
paroxysmal nocturnal haemoglobinuria

Inflammatory Disorders

inflammation of autoimmune diseases
Crohn's Disease
adult respiratory distress syndrome

thermal injury including burns or frostbite
uveitis
psoriasis
asthma
acute pancreatitis

Post-Ischemic Reperfusion Conditions

myocardial infarction
balloon angioplasty
atherosclerosis (cholesterol-induced) & restenosis
hypertension
post-pump syndrome in cardiopulmonary bypass or renal haemodialysis
renal ischemia
intestinal ischaemia

Infectious Diseases or Sepsis

multiple organ failure
septic shock

Immune Complex Disorders and Autoimmune Diseases

rheumatoid arthritis
systemic lupus erythematosus (SLE)
SLE nephritis
proliferative nephritis
glomerulonephritis
haemolytic anemia
myasthenia gravis

Reproductive Disorders

antibody- or complement-mediated infertility

Wound Healing

The present invention is also directed to a pharmaceutical composition comprising a therapeutically effective amount of a polypeptide, as above, and a pharmaceutically acceptable carrier r excipient.

The present invention also provides a method of treating a disease or disorder associated with inflammation or inappropriate complement activation comprising

administering to a subject in need of such treatment a therapeutically effective amount of a polypeptide of this invention.

In the above methods, the subject is preferably a human.

An effective amount of the polypeptide for the treatment of a disease or disorder is in the dose range of 0.01-100mg/kg; preferably 0.1mg-10mg/kg.

For administration, the polypeptide should be formulated into an appropriate pharmaceutical or therapeutic composition. Such a composition typically contains a therapeutically active amount of the polypeptide and a pharmaceutically acceptable excipient or carrier such as saline, buffered saline, dextrose, or water. Compositions may also comprise specific stabilising agents such as sugars, including mannose and mannitol, and local anaesthetics for injectable compositions, including, for example, lidocaine.

Further provided is the use of a polypeptide of this invention in the manufacture of a medicament for the treatment of a disease or disorder associated with inflammation or inappropriate complement activation.

In order to inhibit complement activation and, at the same time, provide thrombolytic therapy, the present invention provides compositions which further comprise a therapeutically active amount of a thrombolytic agent. An effective amount of a thrombolytic agent is in the dose range of 0.01-10mg/kg; preferably 0.1-5mg/kg. Preferred thrombolytic agents include, but are not limited to, streptokinase, human tissue type plasminogen activator and urokinase molecules and derivatives, fragments or conjugates thereof. The thrombolytic agents may comprise one or more chains that may be fused or reversibly linked to other agents to form hybrid molecules (EP-A-0297882 and EP 155387), such as, for example, urokinase linked to plasmin (EP-A-0152736), a fibrinolytic enzyme linked to a water-soluble polymer (EP-A-0183503). The thrombolytic agents may also comprise muteins of plasminogen activators (EP-A-0207589). In a preferred embodiment, the thrombolytic agent may comprise a reversibly blocked *in vitro* fibrinolytic enzyme as described in U.S. Patent No. 4,285,932. A most preferred enzyme is a p-anisoyl plasminogen-streptokinase activator complex as described in U.S. Patent No. 4,808,405, and marketed by SmithKline Beecham Pharmaceuticals under the Trademark EMINASE (generic name anistreplase, also referred to as APSAC; Monk *et al.*,1987, Drugs 34:25-49).

Routes of administration for the individual or combined therapeutic compositions of the present invention include standard routes, such as, for example, intravenous infusion or bolus injecti n. Active complement inhibitors and thrombolytic agents may be administered together or sequentially, in any order.

The present inventi n also provides a method for treating a thrombotic condition, in particular acute myocardial infarction, in a human or non-human animal.

This method comprises administering to a human or animal in need of this treatment an effective amount of a polypeptide according to this invention and an effective amount of a thrombolytic agent.

Also provided is the use of a polypeptide of this invention and a thrombolytic agent in the manufacture of a medicament for the treatment of a thrombotic condition in a human or animal. Such methods and uses may be carried out as described in WO 91/05047.

This invention further provides a method for treating adult respiratory distress syndrome (ARDS) in a human or non-human animal. This method comprises administering to the patient an effective amount of a polypeptide according to this invention.

The invention also provides a method of delaying hyperacute allograft or hyperacute xenograft rejection in a human or non-human animal which receives a transplant by administering an effective amount of a polypeptide according to this invention. Such administration may be to the patient or by application to the transplant prior to implantation.

The invention yet further provides a method of treating wounds in a human or non-human animal by administering by either topical or parenteral e.g. intravenous routes, an effective amount of a polypeptide according to this invention.

GENERAL METHODS USED IN EXAMPLES

(i) **DNA cleavage**

Cleavage of DNA by restriction endonucleases was carried out according to the manufacturer's instructions using supplied buffers. Double digests were carried out simultaneously if the buffer conditions were suitable for both enzymes. Otherwise double digests were carried out sequentially where enzyme requiring the lowest salt concentration was added first to the digest. Once that digest was complete the salt concentration was altered and the second enzyme added.

(ii) **Production of blunt ended DNA fragments**

The recessed 3' termini of DNA fragments were filled in using the Klenow fragment of DNA polymerase I as described in Sambrook *et al* (1989).

(iii) **DNA purification/ concentration and analysis**

Removal of protein contaminants, nucleosides was with phenol/$CHCl_3$ followed by precipitation with ethanol. DNA was analysed on horizontal agarose gel electrophoresis; both methods are described in Sambrook *et al* (1989).

(iv) **DNA fragment isolati n**

1. **DNA purificati n n DEAE NA45 membranes**

DNA fragments were purified from agarose gels by making an incision in the

agarose above and just below the required DNA fragment. NA45 membranes from Schleicher & Schuell (Anderman, Great Britain) that had been soaked in TE (10 mM Tris pH 8.0, 1 mM EDTA) were inserted into the incisions and current reapplied to the gel until the DNA fragment was trapped on the lower membrane; higher molecular weight DNA was trapped on the upper membrane. The lower membrane was removed from the gel and the DNA eluted into 0.05 M arginine/1 M NaCl at 70°C for 2 hours. The DNA was then concentrated by ethanol precipitation as described in Sambrook *et al* (1989).

2. **Electroelution**

DNA fragments were excised from agarose gels and DNA extracted by electroelution using the Unidirectional Electroeluter (IBI Ltd.,Cambridge, England) according to the manufacturer's instructions.

3. **Gel purification**

DNA fragments were excised from agarose gels and DNA extracted using the QIAEX gel extraction kit according to the manufacturers instructions (QIAGEN Inc., USA).

(v) **Plasmid preparation**

Large scale plasmid preparation of plasmid DNA was carried out using CsCl as described in Sambrook *et al* (1989) or using Magic Maxipreps (Promega Corporation, Madison, USA) according to the manufacturers instructions. Mini-plasmid preparations were carried out using either the alkaline lysis method described in Sambrook *et al* (1989) or Magic Minipreps (Promega Corporation, Madison, USA) according to the manufacturer's instructions.

(vi) **Introduction of plasmid DNA into *E. coli***

1. Plasmids were transformed into *E. coli* HB101 or *E. coli* BL21 (DE3) (Studier and Moffat, 1986) that had been made competent using calcium chloride as described in Sambrook *et al* (1989).

2. Alternatively plasmids were introduced into *E.coli* DH1 (Low,1968) or *E.coli* BL21 (DE3) by electroporation using the Gene Pulsar and Pulse Controller of Bio-Rad (Bio-Rad Laboratories, Richmond, California,USA) according to the manufacturer's instructions.

(vii) **Kinasing of oligonucleotides**

Oligonucleotides or annealed oligonucleotides possessing 5' overhangs were kinased using T_4 polynucleotide kinase as described in Sambrook *et al* (1989).

(viii) **Annealing and ligation of oligonucleotides**

Oligonucleotides were annealed together by mixing generally equimolar concentrations of the complementary olig nucleotides in 10 mM Tris pH 8.5, 5 mM $MgCl_2$ and placing at 100°C for 5 minutes and then cooling very slowly to room

temperature. Annealed oligonucleotides with sticky ends were ligated to vector or other oligonucleotides containing complementary sticky ends using T_4 DNA ligase as described in Sambrook *et al* (1989).

(ix) **PCR (Polymerase Chain Reaction) amplification of DNA**

DNA fragments from ligation reactions or DNA fragments excised and purified from agarose gels were amplified by PCR from two primers complementary to the 5' ends of the DNA fragment. Approximately 0.1 - 1µg of ligation reaction or the purified DNA from the agarose gel was mixed in 10 mM Tris pH 8.3 (at 25°C), 50 mM KCl, 0.1% gelatin; $MgCl_2$ concentrations were varied from 1.5 mM to 6 mM to find a suitable concentration for each reaction. Both primers were added to a final concentration of 2 µM; each dNTP was added to a final concentration of 0.2 mM. The final reaction volume was either 75 µl or 100 µl, which was overlayed with mineral oil to prevent evaporation. Thermal cycling was then started on a thermal cycler eg. Hybaid Thermal reactor, and a typical example of conditions used was 94°C 7 mins, 45°C 2 mins, hold at 45°C for less than 5 min., and then add 5 units of Taq DNA polymerase (purchased from a commercial source, e.g. Gibco). The DNA fragment was amplified by cycling the temperature at 72°C 2 mins, 94°C 1 min and 45°C 2 min a total of 35 times.

(x) **DNA sequencing using the double stranded method**

Sequencing was carried out using "Sequenase™" (United States Biochemical Corporation) essentially as described in the manufacturer's instructions.

(xi) **DNA sequence analysis and manipulation**

Analysis of sequences were carried out on a digital VAX computer using the GCG package of programmes as described in Devereux *et al* (1984).

(xii) **Production of oligonucleotides**

1. Oligonucleotides were synthesised using a Gene Assembler Plus (Pharmacia LKB Biotechnology, Milton Keynes, England) or a 381A Synthesiser (Applied BioSystems) according to the manufacturer's instructions.

2. Oligonucleotide purification was carried out either using MonoQ as recommended by Pharmacia or by UV shadowing where recovery of synthetic oligonucleotides was by electrophoresis through a denaturing polyacrylamide gel. The oligonucleotides were loaded onto a 12% acrylamide/7M urea gel and run at 1500V until the oligonucleotide had migrated approximately two thirds of the length of the gel. The DNA was visualised using a hand-held, long-wavelength ultraviolet lamp; and the DNA bands excised. The oligonucleotide was rec vered using Sep-Pak C18 reverse phase columns (Waters) as described in Sambrook et al (1989).

(xiii) **Sodium d decyl sulphate polyacrylamide gel electrophoresis (SDS PAGE)**

SDS PAGE was carried out generally using the Novex system (British Biotechnology) according to the manufacturer's instructions. Prepacked gels of acrylamide concentrations 14%, 16%, 4 - 20% or 10 - 27% were the ones most frequently used. Samples for electrophoresis, including protein molecular weight standards (LMW Kit, Pharmacia) were usually diluted in 1%(w/v)SDS - containing buffer (with or without 5%(v/v) 2-mercaptoethanol), and left at room temperature for about 0.5 to 1 h before application to the gel.

(xiv) **Alteration of codon usage**

The non random use of synonymous codons has been demonstrated in *E. coli* and there is some evidence to support the belief that protein production from genes containing non-optimal or minor codons (particularly at the 5' end of the gene) is less efficient than that from genes with no such codons (e.g. Chen an Inouye, 1990). A codon usage table compiled from genes highly expressed in *E. coli* (supplied as part of the GCG sequence analysis software package, Devereux *et al*, (1984)) was used to determine the optimal codons for expression in *E. coli*. All of the first 30 codons of all constructs (where compatible with restriction enzyme sites) were optimised for high level expression. The codons for the seven amino acids: arg, gly, ile, leu, pro, ser, ala were optimised (where compatible with restriction enzyme sites) throughout the coding sequence.

(xv) **Construction of vector pBROC413**

The plasmid pT7-7 (Tabor, 1990) contains DNA corresponding to nucleotides 2065-4362 of pBR322 and like pBR322 can be mobilized by a conjugative plasmid in the presence of a third plasmid ColK. A mobility protein encoded by ColK acts on the *nic* site at nucleotide 2254 of pBR322 initiating mobilization from this point. pT7-7 was digested with *Lsp*I and *Bgl*II and the protruding 5' ends filled in with the Klenow fragment of DNA Polymerase I. The plasmid DNA fragment was purified by agarose gel electrophoresis, the blunt ends ligated together and transformed into *E. coli* DH1 by electroporation. The resultant plasmid pBROC413 (Fig.1) was identified by restriction enzyme analysis of plasmid DNA.

The deletion in pBROC413 from the *Lsp*I site immediately upstream of the ø10 promoter to the *Bgl*II site at nucleotide 434 of pT7-7 deletes the DNA corresponding to nucleotides 2065-2297 of pBR322. The *nic* site and adjacent sequences are therefore deleted making pBROC413 non mobilizable.

(xvi) **Haemolytic assay**

The anti-haemolytic activity of polypeptides was assessed by measuring the inhibition of complement mediated lysis of sheep erythrocytes sensitised with rabbit antibodies (obtained from Diamedix Corporation, Miami, USA). Human serum diluted 1:125 in 0.1 M Hepes/ 0.15 M NaCl pH 7.4 buffer was the source of

complement and was prepared from a pool of volunteers essentially as described in (Dacie & Lewis, 1975). Briefly, blood was warmed to 37°C for 5 minutes, the clot removed and the remaining serum clarified by centrifugation. The serum fraction was split into small aliquots and stored at -196°C. Aliquots were thawed as required and diluted in the Hepes buffer immediately before use.

Inhibition of complement-mediated lysis of sensitised sheep erythrocytes was measured using a standard haemolytic assay using a v-bottom microtitre plate format as follows.

50 μl of a range of concentrations (0.01-100μg/ml but typically 0.05 - 25 μg/ml) of test protein diluted in Hepes buffer were incubated with 50 μl of the diluted serum for 15 minutes at 37°C. 100 μl of prewarmed sensitised sheep erythrocytes were added for 1 hour at 37°C in a final reaction volume of 200 μl. Samples were spun at 300g at 4°C for 15 minutes before transferring 150 μl of supernatant to flat bottomed microtitre plates and determining the absorption at 410 nm, which reflects the amount of lysis in each test solution. Maximum lysis was determined by incubating serum with erythrocytes in the absence of any inhibitor from which the proportion of background lysis had been subtracted (determined by incubating erythrocytes with buffer). The background lysis by inhibitor was assessed by incubating inhibitor with erythrocytes and then subtracting that from test samples. Inhibition was expressed as a fraction of the total cell lysis such that IH50 represents the concentration of inhibitor required to give 50% inhibition of lysis.

(xvii) **C3a RIA Assay**

Activation of complement pathways can be followed by measuring the release of the anaphylatoxin, C3a and its breakdown product C3a des Arg. Both products can be measured using a competitive radio-immuno assay purchased from Amersham International plc, U. K., (human complement C3a des Arg [^{125}I]assay, code RPA 518).

(a) **Alternative pathway activation by Zymosan A**

The alternative pathway of complement was activated with zymosan A, a complex carbohydrate from yeast (Sigma, catalogue number Z-4250). Zymosan A was made 50 mg/ml in Hepes buffer (0.1M Hepes/0.15M NaCl pH 7.4) or in PBS (50 mM sodium phosphate/0.1 M NaCl pH 7.4) and vortexed until a fine suspension had formed. Serum (prepared as described for the haemolytic assay; Method xvi) was preincubated with different concentrations of complement inhibitor diluted in Hepes buffer for 15 mins at 37°C using the volumes given below. Zymosan A was then vortexed for a few seconds each time before addition to the samples after which samples were incubated for a further 30 mins at 37°C. The zymosan A was then spun down at approximately 11,000g for 30 seconds at ambient temperature.

Typically 100 µl of supernatant were added to an equal volume of precipitating solution provided in the kit and the subsequent supernatant assayed as described in the technical bulletin supplied by Amersham with the C3a des Arg assay RIA kit. Each sample was assayed in duplicate and useful dilutions of the supernatant, to ensure that sample counts were on the standard curve, were found to be 1/50 - 1/100. EDTA or Futhan were not used in any solutions or tubes as suggested in the technical bulletin.

Each sample was counted for 1 minute on an LKB-Wallac 1272 Clinigamma. Data was processed using the RiaCalc program for RIA assays as supplied with the Clinigamma. The data was computed essentially as described in the Amersham technical bulletin with the standrd curve constructed by non-linear regression fit to the data.

The miniaturised assay was carried out essentially as described above but using smaller total volumes for the activation of serum.

Volumes of samples added

	serum	inhibitor	Zymosan A
Normal Assay	79 µl	20 µl	21 µl
Miniaturised Assay	26.3 µl	6.7 µl	7 µl

In the minaturised assay, after activation, typically 25 µl of the sample were precipitated. The assay kit reagent additions were reduced from 50 µl to 10 µl which enabled the assay to be carried out in a U-bottom microtitre plate containing separate detachable wells. The assay was then carried out as described in the technical bulletin using the adjusted volumes until the last dilution in isotonic saline. In this instance 200 µl of saline were added and the plate spun at approximately 2500g for 12 mins at 4°C. The supernatants from each well were carefully removed by aspiration and the precipitate was washed with a further 300 µl of isotonic saline. The plate was then spun again at about 2500g for 5 mins, 4°C and the supernatant was discarded. Wells were then counted for 10 mins each on the Clinigamma. The data was processed as above.

To determine the % inhibition of maximum activation at each inhibitor concentration, a number of controls were carried out with each experiment. These included maximum activation (A) *i.e.* serum + zymosan A only, background activation (B) *i.e.* serum + buffer only, and background activation in the presence of inhibitor (C) *i.e.* serum + inhibitor only. The background activation was generally subtracted from the maximum activation. Similarly the background activation in the presence of inhibitor was subtracted from the value of activated serum in the presence of inhibitor. These values could then be used to determine the % inhibition at each

inhibitor concentration. using the following formula:

$$1 - \frac{(D - C)}{(A - B)} X\ 100$$

where D is the value of activation of serum in the presence of inhibitor and zymosan A. The IC50 is defined as the concentration of inhibitor required to reduce maximum activation by 50%.

(b) **Classical pathway activation by heat aggregated IgG**

Activation of the classical pathway by IgG was performed as follows. Human γ-globulin (Sigma, catalogue number G-4386) was made 14 mg/ml in 0.1 M Hepes/0.15 M NaCl pH 7.4 and heated at 60°C for 1 hour. Samples of heat aggregated IgG were then stored as small aliquots at -80°C until required. Serum was activated using heat aggregated IgG using the same volumes as described for the zymosan A normal or miniaturised assay. Preincubation of inhibitor with serum was for 15 mins at 37°C followed by addition of the heat aggregated IgG. Incubation was continued for a further 45 mins at 37°C. The samples were then assayed directly for C3a levels using either the normal or miniaturised assay.

(xviii) **C5a RIA Assay**

Activation of complement pathways can be followed by measuring the release of the anaphylatoxin C5a and its breakdown product C5a des Arg. Both products can be measured using a competitive radio-immuno assay purchased from Amersham International plc, U. K., (human complement C5a des Arg [^{125}I]assay, code RPA 520).

The alternative pathway of complement was activated with zymosan A, as described for the C3a RIA assay (Method (xvii)). The assay was carried out in the miniaturised form as described for the C3a assay using the reagents provided in the C5a des-Arg RIA kit.

References in the Examples to amino acid numbering relate to the corresponding residues of mature CR1 protein.

Example 1 Construction of plasmid pDB1010-D11 encoding SCR 1 + 2

General points

A DNA sequence for SCR 1 + 2 corresponding to amino acid 1 and ending at amino acid 124 of mature human complement receptor 1 was designed such that the 5' end of the gene contained an *Nde*I site. This site comprises an ATG codon to give the initiating methionine required for the start of mRNA translation and places the gene an ptimum distance from the Shine-Dalgarno ribosome binding sequence of pBROC413. The 3' end of the gene finished on two stop codons followed by a

*Hind*III site.

Restriction endonucleases that do not cut pBROC413 and that were commercially available were identified. The sequences of the restriction sites recognised by the endonucleases were translated into all three reading frames. The sites that contained rarely used codons for *E. coli* expression were discarded. The remaining sites were matched with the DNA coding for SCR 1 + 2. If the restriction site could be fitted into the DNA sequence so as to preserve the coding sequence and not add a rarely used codon, the DNA sequence was altered to include this restriction site. 10 unique restriction sites were so identified and incorporated. To enable intracellular expression of protein in *E. coli*, an ATG codon was added to the 5' end of the gene immediately preceding the codon for the first amino acid of mature CR-1. The codon ATG is part of the *Nde*I restriction site which can be used for cloning into vectors such as pBROC413. The codon corresponding to proline 124 of mature CR-1 has been changed to one encoding glutamine, which also encompasses an *Eco*RI site.

(a) **Construction of plasmid**

Oligonucleotides coding for SCR 1 + 2 (Table 1; 1 - 8) were synthesised as 4 complementary pairs of 87 - 101 mers that could be ligated in a unique fashion via complementary 8 bp overhangs between the pairs of oligonucleotides. The four complementary pairs of oligonucleotides were designated Pair A (oligos 1+2), Pair B (oligos 3+4), Pair C (oligos 5+6) and Pair D (oligos 7+8). Pair A which corresponded to the 5' end of the gene contained an *Nde*I restriction site overhang and Pair D contained a *Hind*III restriction site overhang at the 3' end. All oligonucleotides apart from 1 and 2 were purified on Pharmacia Mono Q columns prior to use. Oligonucleotide 2 of pair A and oligonucleotide 7 of pair D were kinased before annealing with their unkinased complementary oligonucleotides 1 and 8 respectively. Oligonucleotides pairs B and C were annealed first and then kinased. The kinased oligonuleotide pairs were ligated Pair A (approx. 0.1µg) to Pair B (approx. 0.2µg) and Pair C (approx. 2µg) to Pair D (approx. 4µg). The ligated oligonucleotides (A+B) were in turn ligated to (C+D) to form the gene coding for SCR 1 + 2.

The DNA coding for SCR 1 + 2 was amplified by PCR using two oligonucleotides (Table 1; 15 and 16) complementary to the two strands of DNA. Both oligonucleotides contained 5' unmatched ends that contained 6 bp of random sequence followed by the sequence f either *Nde*I or *Hind*III restriction sites followed by 18 bp complementary to the gene. Following PCR, a band of approximately 400 bp was visualised on horizontal agarose gel electrophoresis, which was excised and purified on DEAE NA45 membranes. The DNA was then cut with *Nde*I and *Hind*III

before ligating into pBROC413 that had been cut with the same enzymes. The vector was transformed into *E. coli* HB101 made competent with calcium chloride. Mini-plasmid preparations were made and the plasmid DNA was analysed by digestion with *Nde*I and *Hin*dIII. Plasmids containing the correct sized insert, were further subjected to restriction mapping with *Eco*RI, *Hpa*I, *Kpn*I and *Sma*I. The plasmids that displayed the correct restriction maps were analysed by DNA sequencing of both strands across the gene coding for SCR 1 + 2. Plasmid pDB1010-D11 was identified as having the correct sequence across the gene coding for SCR 1 + 2.

Example 2 Construction, expression, purification, folding and formulation of MQ1 -> K196 of CR-1 (SCR 1+2+3)

General Points

The DNA coding for SCR 1+2+3 was constructed by ligating DNA coding for SCR 1+2 (Example 1a) to DNA encoding SCR 3.

General points relating to SCR 3 are presented in Example 9.

The SCR 3 coding unit corresponding to amino acid 122 and ending at amino acid 196 of mature CR1, was designed such that 5' end of the unit contained the *Eco*RI site at the junction of SCR's 2 & 3 as well as an *Nde*I site 5' to the *Eco*RI site. The 3' end of the unit finished on two stop codons followed by a *Hin*dIII site. The plasmids containing the SCR 3 coding unit and the SCR 1+2 coding unit were digested with *Eco*RI and *Hin*dIII. The SCR 3 coding unit was isolated and inserted downstream of the SCR 1+2 coding unit in the *Eco*RI/*Hin*dIII-cut SCR 1+2 - containing plasmid, to give a plasmid containing the SCR 1+2+3 c[illegible]g unit, which corresponds to amino acids 1 to 196 of mature CR1. The addition of the SCR 3 coding unit through the *Eco*RI site, converts the codon corresponding to a glutamine at position 124 back to the authentic amino acid (proline) that is found in CR1.

(a) **Construction of plasmid pDB1013-5-4 encoding SCR 1 + 2 + 3**

Three pairs of oligonucleotides (Table 1; 9 - 14) encompassing the SCR[illegible] coding sequence were synthesised. The oligonucleotides were firs[illegible] [illegible]aled a[illegible] (9, 10; 11,12; 13,14) and the middle pair kinased thus allowing [illegible] three pa[illegible] [illegible]e ligated together via 8 base pair overlapping sequences. The 5' end of this [illegible] molecule was designed to be complementary to *Nde*I digested DN[illegible] and th[illegible] [illegible] to *Hin*dIII digested DNA. This enabled the trimer to be cloned into *Nde*I/*Hin*dIII digested pBROC413 generating pBROC435 (Fig.2). The identity of pBROC435 was checked by restriction enzyme analysis and confirmed by DNA sequencing.

Plasmid DNA from pBROC435 and pDB1010-D11 (Example 1) were both cut with *Eco*RI and *Hin*dIII; the *Eco*RI/*Hin*dIII band of pBROC435 coding for SCR 3

was purified on an DEAE NA45 membrane as was the cut vector pDB1010-D11. The SCR 3 coding unit was then ligated into pDB1010-D11 to generate pDB1013-5 which was then transformed into calcium chloride competent *E. coli* HB101. The resulting colonies were analysed by mini-plasmid preparation of DNA followed by restriction mapping. One of the colonies, termed pDB1013-5-4 (Fig.2), contained the SCR 1+2+3 coding unit. This plasmid was then analysed for expression of the gene product.

(b) **Expression of SCR 1 + 2 + 3**

pDB1013-5-4 was transformed into calcium chloride competent *E.coli* BL21(DE3) and resulting colonies analysed by restriction digestion of mini-plasmid DNA preparations. Single colonies were inoculated into universals containing 10 ml of L broth or NCYZM medium and 50 μg/ml ampicillin and allowed to grow overnight at 37°C, 220 r.p.m. The overnight cultures (typically 5ml) were used to inoculate each of 2 L conical flasks containing 500 mls of NCYZM medium, 150 μg/ml ampicillin; cultures were grown at 37°C, 220 r.p.m. until A_{600} was 0.5 absorbance units. Cultures were induced with 1 mM isopropylthioβ-D-galactoside (IPTG) and allowed to grow a further 3 hours under the same conditions. The cultures were centrifuged (approx. 8000g/10 min) and the supernatants discarded. The cell pellets were stored at -40°C. L broth was 1% (w/v) Bactotryptone, 0.5% (w/v) Bactoyeastextract, 0.5% (w/v) NaCl. NCYZM media was L-broth containing 0.1% (w/v) casamino acids and 0.2%(w/v) $MgSO_4.7H_2O$, pH 7.0.

(c) **Isolation of solubilised inclusion bodies**

Frozen cell pellet of *E.coli* BL21 DE3 (pDB1013-5-4) (1 litre culture) prepared in a similar way to that described in Example 2b. was allowed to thaw at 4°C for 2 h and was then resuspended in 50 mM Tris/50 mM NaCl/1 mM EDTA/0.1 mM PMSF pH 8.0 (33 ml). The suspension was transferred to a 100 ml glass beaker and sonicated (Heat Systems - Ultrasonics W380; 70 Watts, 50 x 50% pulse, pulse time = 5 sec.). The sonicate was immediately centrifuged (6000g/4°C/10 min) and the supernatant was discarded. The pellet, containing the inclusion bodies, was resuspended in 20 mM Tris/8 M urea/50 mM 2-mercaptoethanol/1 mM EDTA/0.1 mM PMSF pH 8.5 (100 ml) and left static at room temperature (approx. 23°C) for 1 h. The resulting solution was centrifuged (approx. 2000 g at 4°C for 10 min) to remove any material that had failed to solubilise. The supernatant of this spin was retained at -40°C as the solubilised inclusion body product.

(d) **Purificati n of SCR 1 + 2 + 3 fr m the solubilised inclusion body**

A column (i.d., 16 mm; h, 10 mm) of S-Sepharose Fast Flow was prepared and connected into an FPLC (Pharmacia) system. The column was equilibrated with 20 mM Tris/8M urea/1 mM EDTA/50 mM 2-mercaptoethanol pH 8.5. 10 ml of

thawed, solubilised inclusion body, prepared as described in Example 2c, was applied to the column and washed through with equilibration buffer. The column was then developed with a linear gradient to 1M NaCl (in equilibration buffer) followed by rinses with 1M NaCl and 2M NaCl (also in equilibration buffer). All the chromatography was at 1.0 ml min $^{-1}$ and at room temperature.

Analysis by SDS PAGE/protein staining of the fractions collected during the chromatography indicated that virtually all the SCR 1 + 2 + 3 polypeptide had absorbed to the column and had been dissociated by the 1M NaCl - containing buffer. The appropriate fractions were stored at -40°C.

Subsequent assay for protein content of the peak fraction using the Bradford protein assay and a bovine serum albumin standard showed it contained 2.8 mg protein.

(e) **Folding**

S-Sepharose - purified SCR 1 + 2 + 3 that had been purified in a similar way to that described in Example 2d and stored at -40°C was thawed and 0.4 ml buffer-exchanged into 0.05 M formic acid using Sephaex G25 (P10). The absorbance at 280 nm of the buffer-exchanged solution was determined as 0.52, and, using ϵ = 34000 and appropriate correction factors for dilution, the protein concentration of the original preparation (prior to buffer-exchange) was calculated to be 0.6 mg/ml.

Based on this figure, 1.7 ml S-Sepharose - purified protein was diluted with 0.85 ml 20 mM Tris/8M urea/50 mM 2-mercaptoethanol/1M NaCl pH 8.5 to give a 0.4 mg/ml solution, on which the folding was carried out.

Folding was effected by a series of dilutions, using cold diluent at all times.

At t=0 h, 0.8 ml SCR1 + 2 + 3 (0.4 mg/ml) was added to 0.8 ml 20 mM Tris/1M urea/5 mM EDTA/3 mM 2-mercaptoethanol pH 8.0 ('diluent') in a 30 ml polystyrene universal container. The solution was mixed thoroughly by gentle swirling and left static, capped, in a cold room (approx. 2 to 3°C).

At 1 h, 1.6 ml diluent was added and mixed.

At 2 h, 3.2 ml diluent was added and mixed.

At 4 h, 6.4 ml diluent was added and mixed.

The solution was left a further 20 h in the cold room, then ultrafiltered (YM5, Amicon Ltd) to approx. 1.4 ml. This was buffer-exchanged into 0.1M NH_4HCO_3 (2.5 ml) using Sephadex G25 (PD 10) in the cold room. The eluate was aliquoted and was stored at -40°C or lyophilised.

The product containing SCR 1 + 2 + 3 was analysed by SDS PAGE, followed by protein staining. In both non-reduced and reduced (with 2-mercaptoethanol 5% (v/v)) samples there was a single major band. The molecular weight of the reduced band, compared to reduced protein standards f known M_r, was approx. 24,000. The

non-reduced protein (band) had a slightly faster mobility than the reduced protein (band).

The product was analysed in a functional haemolytic assay utilising antibody-sensitized sheep erythrocytes (Method (xvi)). The product showed concentration-dependent inhibition of the complement - mediated lysis of the erythrocytes with an IH50 around 0.4 µg/ml.

(f) **Folding**

Preparation, folding, processing and analysis were carried out exactly as described in Example 2e except

(1) the diluent for the folding was 20 mM ethanolamine (pH 10.0)

(2) the folded solution was ultrafiltered to a final volume of 1.55 ml, and

(3) the IH50 figure was determined as about 0.6 µg/ml.

(4) the recovery of product was approx. 100 per cent.

(g) **Determination of N-terminal sequence of SCR 1 + 2 + 3**

1ml samples of growing *E.coli* BL21 (DE3) containing plasmid pDB1013-5-4 were removed 3 hours post-induction with 1 mM IPTG as described in Example 2b. These samples were spun in an eppendorf centrifuge and the resultant pellets each resuspended in 200µl of reducing buffer (100 mM Tris pH6.8/200 mM dithiothreitol/4% (^{w}/v) SDS/2% (^{w}/v) bromophenol blue and 20% (^{v}/v) glycerol and boiled for 5 minutes. 25µl samples were applied to a 14% polyacrylamide gel. When the electrophoresis was complete the proteins were transferred to a ProBlott membrane (Applied Biosystems) using a Sartoblot electroblotting apparatus (Sartorius) at 0.8 mA/cm^2 for 1 hour 40 mins using CAPS (3-[cyclohexylamino]-1-propanesulphonic acid) transfer buffer. After transfer the ProBlott membrane was stained (0.1% (^{w}/v) Coomassie Blue R-250/40% (^{v}/v) methanol/1% ^{v}/v acetic acid) for 20 seconds and destained using 50% (^{v}/v) methanol. A band corresponding to a M_r approx 23,000 protein was excised and the N-terminal sequence determined using a Blott cartridge in an Applied Biosystems 477A Protein Sequencer.

The sequence of the first 20 amino acids was found to agree with the predicted sequence except that residue 3 could not be identified by the sequencing protocol used.

Example 3 Expression and purification of SCR 1 + 2 + 3 from a fermentation vessel

(a) **Fermentati n of *E. coli* harbouring the plasmid pDB1013-5-4**

E coli BL21 (DE3) : pDB1013-5-4 was recovered from storage in liquid nitrogen by thawing a vial containing 1 ml f the culture and this was used to inoculate 100 ml of seed medium (NCYZM) containing ampicillin at 75 µg/ml. The

primary and secondary seed stage fermentations were carried out in plain 500 ml shake flasks batched with 100 ml aliquots of NCYZM medium. The primary and secondary seed fermentation conditions were as follows : 37°C, 230 rpm on an orbital shaker with a 50 mm throw. The primary seed incubation time was 2 hours. The primary seed culture was used to inoculate secondary seed fermentation medium at 0.1% (v/v). The secondary seed was incubated for 14.5 hours.

Two 15 litre Biolafitte fermenters were each batched with 10 litres of NCYZM medium and 0.01% (v/v) Dow Corning DC1510 antifoam. The vessels plus media were sterilised using steam to 121°C for 45 minutes. Ampicillin sterilised by microfiltration (0.2 µm) was added aseptically to the vessel media to a final concentration of 150 µg/ml. The fermenters were inoculated at a level of 3% (v/v) from pooled secondary seed culture. The final stage incubation conditions were 37°C, agitator 400 rpm, airflow 5 l/min (0.5 vvm). The final stage fermentations were sampled aseptically pre-inoculation, at 0 hours and thence every half hour. The samples were monitored for increases in optical density (600 nm). When the OD600 was ≥ 0.5, IPTG was added to give a final concentration of 1 mM. The fermentations were incubated for a further 3 hours.

The cells were recovered by centrifugation using 7000 g for 25 minutes. The total cell yield (wet weight) was 49.8 grammes.

(b) **Inclusion body isolation.**

Inclusion bodies from 23 g (wet weight) cell pellet were isolated and solubilised essentially as described in Example 2.

(c) **Purification of denatured SCR 1 + 2 + 3**

The volume of solubilised inclusion body from Example 3b was approx. 800 ml. To this viscous solution was added SP-Sepharose FF (100 ml gel bed, water washed and suction dried). The mixture was swirled vigorously and left static for 1h at room temperature. The supernatant was decanted and stored at -40°C. The remaining slurry was resuspended to a uniform suspension and poured into a glass jacket (i.d., 41.5 mm) and allowed to settle into a packed bed. This packed bed was connected into a low pressure chromatography system at 4°C and equilibrated with 0.02M Tris/8M urea/0.05M 2-mercaptoethanol/0.001 M EDTA pH 8.5. When the A_{280} of the eluate had minimised, the buffer was changed (step-wise) to the equilibration buffer additionally containing 1 M NaCl. A single A_{280} peak was eluted, in a volume of 90 ml (equivalent to approx. 1 Vt). The solution was clear and colourless and was estimated, by A_{280} determination of a buffer-exchanged sample (using an $\varepsilon = 25{,}000$), to contain about 300 mg target protein. By SDS PAGE followed by prot in stain the target protein was the major band present. The 90 ml product was stored at -40°C.

(d) **Folding and further purificati n.**

18 ml of the above product (nominal 60 mg) was diluted with 12 ml 0.02 M Tris/8M urea/1 M NaCl/0.05M 2-mercaptoethanol pH 8.5. The product (30 ml) was added as 5 ml aliquots at 1 min intervals to 930 ml freshly prepared, cold 0.02 M ethanolamine/0.001 M EDTA, with swirling, and left static for 1 h/4°C. Then reduced glutathione was added to 1 mM (by addition of 9.6 ml 0.1M GSH) and oxidised glutathione was added to 0.5 mM (by addition of 9.6 ml 0.05M GSSG). The solution was clear and was left static in the cold for approx. 70 h. The solution was then ultrafiltered using a YM10 membrane to a final retentate volume of about 10 ml; this retentate was cloudy. It was mixed with 90 ml 0.1 M NaH_2PO_4/1 M $(NH_4)_2SO_4$ pH 7.0 (Buffer A) at room temperature and then centrifuged at 4000 rpm for 20 min. The supernatant was decanted and SCR 1 + 2 + 3 protein isolated by chromatography of the supernatant on Butyl Toyopearl 650 S.

The column of Butyl Toyopearl (Vt ~ 12 ml) was equilibrated with Buffer A. The 100 ml supernatant was applied to the column and the column washed with Buffer A. It was then developed with a linear gradient of 100% Buffer A to 100% 0.1 M NaH_2PO_4 pH 7.0. All the chromatography was at room temperature at approx. 30 cmh^{-1}.

A major A_{280} peak was eluted during the gradient. Fractions spanning the peak were analysed by SDS PAGE followed by protein stain. The most concentrated fractions of the peak contained essentially pure SCR 1 + 2 + 3 and were active in the haemolytic assay (Method (xvi)) (IH_{50} ~ 0.3 µg/ml). They were stored at -40°C.

Example 4 Formulation of Butyl Toyopearl purified SCR 1 + 2 + 3.

Batches of SCR 1 + 2 + 3 that had been expressed, folded and purified in similar ways to batches described in Examples 2 and 3 and further purified by ammonium sulphate treatment and Butyl Toyopearl chromatography essentially as described in Example 3d were formulated into a useable product as follows.

Three such Butyl Toyopearl products were pooled to give a volume of about 31 ml. All 31 ml were buffer-exchanged into 0.05 M formic acid (prepared using 0.2 µm-filtered 'MilliQ' water) using a column of Sephadex G25. All the chromatography was at 50 cmh^{-1} at 4°C. The eluate from the column was monitored at 280 nm and the Vo fraction was collected as a single fraction. The bulk of this fraction was lyophilised in aliquots.

Analysis of the Vo pool prior to lyophilisation by both SDS PAGE/stain and C8 reverse phase HPLC showed it to be essentially pure target protein. The pool demonstrated anti-haemolytic activity (IH_{50} approx. 0.3 µg/ml) and the endotoxin content was low (< 1ng/mg).

One of the lyophilised aliquots was shown to be soluble at 10 mg ml^{-1} in phosphate-buffered saline and showed complement inhibitory activity in the haemolytic assay (Method xvi); the IH50 was 0.3μg/ml.

Another of the lyophilised aliquots was examined to determine the disulphide bridge pattern. All six correct (as predicted on the basis of a consensus SCR motif) disulphides were detected.

Example 5 Effect of SCR 1+2+3 on IgG-mediated activation of the classical pathway of complement, as measured by C3a release

Inhibition of heat aggregated IgG activated serum was carried out as described in Method (xvii). Heat aggregated IgG activates the classical pathway of complement. Different concentrations (typically 4 - 1000 μg/ml) of inhibitor were incubated with serum in the presence of heat aggregated IgG and the % inhibition of activation at each concentration was determined. The IC50 of SCR 1+2+3 was determined as approximately 22 μg/ml indicating that SCR 1+2+3 can inhibit the classical pathway of complement.

Example 6 Effect of SCR 1+2+3 on zymosan A-mediated activation of the alternative pathway of complement, as measured by following C3a release.

Inhibition of zymosan A activated serum was carried out as described in Method (xvii). Different concentrations of SCR 1+2+3 (typically in the range 1 - 1000 μg/ml) were incubated with serum in the presence of zymosan A and the % inhibition of activation at each concentration was determined. From several different experiments the IC50 was determined as 20 - 40 μg/ml indicating that SCR 1+2+3 can inhibit the alternative pathway of complement.

Example 7 Effect of SCR 1+2+3 on zymosan A-mediated activation of the alternative pathway of complement, as measured by C5a release.

Inhibition of zymosan A activated serum was carried out as described in Method (xvii) and assayed as described in Method (xviii). Different concentrations of SCR 1+2+3 (typically in the range 4 - 700 μg/ml) were incubated with serum in the presence of zymosan A and the % inhibition of activation at each concentration was determined. From several different experiments the IC50 was determined as approximately 20 - 30 μg/ml, indicating that SCR 1 + 2 + 3 can inhibit the alternative pathway of complement.

Example 8 End t xin content determination of purified, folded and formulated SCR 1 + 2 + 3

A batch f final product SCR 1 + 2 + 3 was prepared essentially as described in Example 4 above and was measured for endotoxin content using a method based on the gel-clot reaction of limulus amoebocyte lysate (LAL) (Atlas Bioscan Ltd.). The sensitivity of the assay was 0.125 EU/ml and this was checked by titration against a doubling dilution series prepared from standard *E. coli* endotoxin supplied with the LAL kit.

10-fold dilutions of ~ 1.3 mg/ml SCR 1 + 2 + 3 protein stock were tested in quadruplicate for their effect on LAL by adding 10μl of sample to 10μl LAL. After 1h at 37°C the mixtures were tested for either clotting or remaining liquid. (Solutions that contain at least 0.125 EU endotoxin will clot this LAL preparation.) After taking into account the results of simultaneous tests designed to test for interference, it was concluded that the endotoxin content of the SCR 1 + 2 + 3 protein preparation was < 12.5 EU/ml, equivalent to approx. < 1 ng/mg protein.

Example 9 Expression, folding, purification, and formulation of MR122 -> K196 of CR-1 (SCR 3)

General Points

The sequence for SCR 3 corresponding to amino acid 122 and ending at amino acid 196 of mature human complement receptor 1 was designed such that the 5' end of the gene contained an *Nde*I restriction endonuclease site. This site comprises an ATG start codon to give the initiating methionine required for the start of mRNA translation and allows the placement of the gene an optimum distance from the Shine-Dalgarno ribosome binding site of pBROC413. This codon is followed immediately by the gene coding for SCR 3 starting with arginine 122 of mature human complement receptor 1. The 3' end of the gene finishes with a codon for lysine 196 followed by two stop codons followed by a *Hin*dIII site.

The DNA coding for SCR3 was modified for optimum codon usage in *E. coli* as described in the methods. The gene was also altered to incorporate unique restriction endonuclease sites. This was carried out in the following way. Restriction endonucleases that do not cut pBROC413 and were commercially available were identified. The DNA sequence of these restriction endonuclease sites was then translated into all three reading frames and the codon usage examined. Sites that contained codons that are rarely used by *E. coli* were discarded. The remaining sites were xamined for their translated sequence and if that sequence matched with SCR 3, the restricti n site was incorporated into the sequence.

(a) **Construction of plasmid pBROC435 encoding SCR 3**

The construction of pBROC435 is described in Example 2a

(b) **Expressi n f SCR 3 from pBROC435**

pBROC435 was transformed by electroporation into *E.coli* BL21(DE3) and resulting colonies analysed by restriction digestion of mini-plasmid DNA preparations. Single colonies were inoculated into universals containing 10 ml of L broth or NCYZM medium and 50 - 75 µg/ml ampicillin and allowed to grow overnight at 37°C, 220 r.p.m. Typically 4ml of overnight cultures were used to inoculate each of 2 L conical flasks containing 500 ml of NCYZM medium, 150 µg/ml ampicillin; cultures were grown at 37°C, 230r.p.m. until A_{600} was 0.5 absorbance units. Cultures were induced with 1 mM IPTG and allowed to grow a further 3 hours under the same conditions. The cultures were centrifuged (approx, 8000g/10 min) and the supernatants discarded. The cell pellets were stored at -40°C.

(c) **Isolation of solubilised inclusion bodies**

The frozen cell pellet of *E. coli* (from 3 l growth culture in NCYZM) described in Example 9b was allowed to thaw at room temperature for 2 h and was then resuspended in 50 mM Tris/50mM NaCl/1mM EDTA/0.1 mM PMSF pH 8.0 (90ml). The suspension was transferred to a 200 ml glass beaker and sonicated (Heat Systems - Ultrasonics W380; 70 Watts, 50 x 50% pulse, pulse time = 5 sec.). The sonicate was immediately centrifuged (6000g/4°C/10 min) and the supernatant was discarded. The pellet, containing the inclusion bodies, was resuspended in 20 mM Tris/8 M urea/50 mM 2-mercaptoethanol/1 mM EDTA/0.1mM PMSF pH 8.5 (300ml) with gentle pipetting to mix. After mixing, the solution was left static at room temperature (approx. 23° C) for 1 h. The resulting solution was centrifuged (2000 g at 4° C for 10 min) to remove material that failed to solubilise. The supernatant of this spin was retained at -40° C as the solubilised inclusion body product.

(d) **Purification of SCR3 from the solubilised inclusion body**

A column (i.d., 32 mm; h, 32mm) of Q-Sepharose Fast Flow (Pharmacia) was prepared and equilibrated with 20 mM Tris/8M urea/50 mM 2-mercaptoethanol pH 9.0. 200 ml of thawed, solubilised inclusion body, prepared as in Example 9c, was applied to the column and washed through with equilibration buffer. The column was connected to an FPLC system and developed via a stepwise gradient of 0.1, 1.0, 2.0M NaCl (also in equilibriation buffer). All chromatography was at 2.0 ml min^{-1} and at room temperature.

Analysis by SDS PAGE/protein staining of the fractions collected during the chromatography indicated that virtually all the SCR3 did not bind to the column. Many other proteins had absorbed to the column h wever and had been dissociated by the 0.1M and 1M NaCl - containing buffers. The purity of SCR3 in the

unadsorbed fraction was estimated to be about 80%.

(e) **Folding of SCR3**

Q-Sepharose-purified SCR3 that had been purified as described in Example 9d and stored at -40°C was thawed and was folded by a series of dilutions, using cold diluent. At t=0, 100 ml 20 mM Tris/1M urea/5mM EDTA/3mM 2-mercaptoethanol pH 8.0 (diluent) were added to 100ml SCR3. At this stage the solution was turbid in appearance. The solution was mixed thoroughly by gentle swirling and left static, capped, in a cold room (2-3°C). At 1 h, 200ml diluent was added and mixed, final volume = 400 ml. At 2 h, 400ml diluent was added and mixed, final volume = 800ml. At 4 h 800ml of diluent was added and mixed, final volume = 1.6L. The solution was left for a further 20 h in the cold room. The solution now appeared clear, and it was stored at -40°C in aliquots.

(f) **Formulation of SCR3**

50 ml of SCR3 prepared as in Example 9e were thawed and ultrafiltered to 3.5 ml using a 2000 Da cut-off membrane (Amicon). 2.5 ml of the concentrate was buffer-exchanged into 0.1M NH_4HCO_3 (3.0 ml) using Sephadex G25 (PD 10). Subsequent analysis for protein content using the molar extinction coefficient of 11000 showed this sample contained approx 0.24mg/ml.

Analysis of this material by SDS PAGE/protein staining indicated that the protein was about 80% pure. Samples reduced with 2-mercaptoethanol had a lower electrophoretic mobility suggesting the presence of disulphide bonds in SCR3.

Analysis of this sample in the haemolytic assay (Method (xvi)) showed it had an IH50 of approx. 10 - 20 μg/ml.

(g) **Determination of N-terminal sequence of expressed SCR3**

200μl SCR3 prepared and formulated in 0.1M NH_4HCO_3 as in Example 9f was precipitated with 800μl cold acetone in a cardice/ethanol bath for 60 mins. The sample was then spun in an Eppendorf centrifuge (approx 10,000g/20 mins) and the resultant pellet resuspended with heating in sample buffer containing 5% (v/v) 2-mercaptoethanol. 30 μl samples were electrophoresed on a 4 to 20% SDS-containing polyacrylamide gradient gel. When the electrophoresis was complete the proteins were transferred to a ProBlott membrane (Applied Biosystems) using an electroblotting apparatus at 200mA for 2 h using CAPS in 10% methanol/90%H_2O (v/v)) transfer buffer. After transfer the ProBlott membrane was stained (0.1%(w/v) Coomassie Blue), destained, rinsed and air dried according to the manufacturer's instructions. Sections of the membrane were excised and used for N-terminal sequencing.

The sequence of the first 20 amino acids of the major band was as expected for SCR3 with the exception of residue 5, which could not be identified.

(h) **Preparation, f lding and f rmulation of SCR3**

Preparation and folding were carried out exactly as described in Example 9a-9e. 400 ml of folded SCR3 was ultrafiltered through a 30 KDa cut-off filter (Amicon) at 4° C. Samples of the ultrafiltrate were processed in two ways.

1. 50 ml were ultrafiltered using a 2 KDa cut-off membrane to a final volume of 3.5ml and buffer-exchanged into 0.05 M formic acid (6.7 ml) using Sephadex G25 (PD10) columns. The total amount of SCR3 estimated by the absorbance at 280 nm was 0.6mg. Analysis by SDS PAGE/protein staining indicated that the protein had a purity of about 95%. The sample was freeze-dried and stored at -40°C.

2. 100ml of the ultrafiltrate were adjusted to pH 5.5 with HCl. The sample was applied to a Mono S column (1ml) at 1.5 ml min^{-1} and washed through with equilibration buffer (20mM Tris.HCl pH 5.5). The column was then developed with a step gradient of 0.1, 1.0 and 2.0M NaCl (also in equilibration buffer). All remaining chromatography was at 1.0 ml min^{-1} and at room temperature.

Analysis by SDS PAGE/protein staining of the fractions collected during the chromatography demonstrated that the major band dissociated at 1M NaCl contained SCR3 at about 95% purity.

Example 10 Expression, folding, purification and formulation of MR122-S253 of CR-1 (SCR 3 + 4)

(a) **Construction of plasmid pDB1019 encoding SCR 3 + 4**

The DNA coding for SCR 3 + 4 was constructed from the plasmids pBROC435 (Example 2) and pDB1018-1 (Example 11) which carry the genes coding for SCR 3 and SCR 1 + 2 + 3 +4 respectively. The SCR 4 coding unit was excised from pDB1018-1 and ligated onto the end of the SCR 3 coding unit in pBROC435.

pDB1018-1 was digested with *Spe*I and *Hin*dIII and separated on a 1 % agarose gel. The band which codes for SCR 4 (~ 245 bp) was excised from the gel and purified using the QIAEX extraction kit. Plasmid pBROC435 was also cut with *Spe*I and *Hin*dIII, separated on 1 % agarose, excised from the agarose and purified with the QIAEX kit. The SCR 4 coding DNA was then ligated into the cut pBROC435 plasmid to give pDB1019. This DNA was used to transform *E. coli* HB101 made competent with $CaCl_2$. Transformants were analysed by restriction analysis using *Eco*RI and *Hin*dIII. Clones carrying the correct sized insert were used for expression studies.

(b) **Expression of SCR 3 + 4 from pDB1019-1C**

pDB1019 was transformed int *E. coli* BL21(DE3) made competent with $CaCl_2$ and the resulting colonies were analysed by restriction digestion f mini-

plasmid DNA preparations. Plasmid pDB1019-1C was identified as carrying the correct sized insert. Single colonies of *E. coli* BL21(DE3) carrying pDB1019-1C were inoculated into ten universals containing 10 mls of NCYZM medium and 75 μg/ml ampicillin and allowed to grow overnight at 37°C, 240 r.p.m. The overnight cultures were then used to inoculate eight 2 L conical flasks (5 ml/flask) containing 500 ml of NCYZM medium, 150 μg/ml ampicillin. Cultures were grown at 37°C, 240 r.p.m. until A_{600} was 0.5 absorbance units. At this point cultures were induced with 1 mM IPTG and allowed to grow a further 3 hours under the same conditions. The cultures were centrifuged (approx. 8000g/ 10 mins) and the supernatants were discarded. The cell pellets were stored at -40°C.

(c) Isolation, purification, folding and formulation of SCR 3 + 4

The methods used generally follow those described earlier for the preparation of SCR 1 + 2 + 3.

Isolation of solubilised inclusion bodies from cell pellet derived from 2l culture was carried out as described in Example 2c. The volume of solubilisate was 200 ml.

Some of the contaminating (host) *E. coli* proteins were removed from the preparation by adsorption onto S-Sepharose, either in a batch process or by column chromatography, using systems similar to those described in Example 2d. The protein present in the unadsorbed fractions was shown by SDS PAGE/stain to contain significant amounts of SCR 3 + 4 protein. About half of these fractions were ultrafiltered using a YM1 (Amicon) membrane to approx. 35 to 40 ml. This retentate was estimated to contain about 0.3 mg protein/ml (based on A_{280} determination of a buffer-exchanged sample, using ε= 21,000). 10.5 ml of the retentate was mixed with 325 ml cold 20 mM ethanolamine and left static at 4°C for 1 hour. Then reduced glutathione was added to 1 mM (by addition of 3.4 ml 100mM GSH) and oxidised glutathione was added to 0.5mM (by addition of 3.4ml 50mM GSSG). The solution was mixed and left static at 4°C for ~ 72 h. The solution was clear. The solution was then ultrafiltered using a YM1 membrane to a retentate of 5 ml. The retentate was divided in two and buffer-exchanged into either 20 mM ethanolamine or 50 mM formic acid using Sephadex G25 (PD10 columns).

Analysis of the formic acid SCR 3 + 4 product by reverse phase HPLC and by SDS PAGE followed by protein staining showed only one major protein species (> 90% pure). The protein concentration was estimated to be 0.3 mg/ml using A_{280} determinations. The product was active in the haemolytic assay (Method (xvi)); the IH50 value was approx. 30 μg/ml

Example 11 Construction, expression, folding, purification and formulati n of MQ1-S253 f CR-1 (SCR 1 + 2 + 3 + 4)

General p ints

Two constructs were prepared by making a plasmid encoding SCR 1+2, incorporating SCR3 and finally adding SCR4. The two constructs encoded consensus SCR1 to 4 and the R235H mutation of SCR1 to 4 (Example 12).

A plasmid containing the SCR 1 + 2 + 3 + 4 coding unit was constructed by adding the DNA encoding SCR 4 onto the construct coding for SCR 1 + 2 + 3 (Example 2). For convenience of DNA manipulation, the SCR 4 DNA coding unit was made by synthesising the DNA encoding the last 17 amino acids of SCR 3 followed by the DNA coding for the linker region followed by SCR 4. This DNA started at the *Spe*I site of the SCR 1 + 2 + 3 coding construct which corresponds to T175 of mature CR-1 followed by the DNA coding for the linker region followed by SCR 4 ending on the codon corresponding to S253 followed by two stop codons and a *Hind*III site. As for the previous constructs the DNA encoding SCR 4 was altered for optimised codon usage and restriction sites as previously described in Example 1. This unit of DNA was ligated to the plasmid coding for SCR 1 + 2 + 3 which had been cut with *Spe*I and *Hind*III to give a construct coding for SCR 1 + 2 + 3 + 4.

(a) **Construction of plasmid pDB1018 encoding SCR 1 + 2 + 3 + 4**

Oligonucleotides (Table 1; oligos 21 - 26 coding for SCR4) were synthesised as 3 complementary pairs of 68-90 mers that could be ligated in a unique fashion via complementary 8 bp overhangs between the pairs of oligonucleotides. The 3 complementary pairs of oligonucleotides were designated Pair E (oligos 21, 22), Pair F (oligos 23, 24) and Pair G (oligos 25, 26). Pair E which corresponds to the 5' end of the gene contained a *Spe*I restriction site overhang and Pair G contained a *Hind* III restriction site overhang at the 3' end. All oligonucleotides apart from 22 and 24 were purified by electrophoresis through a denaturing polyacrylamide gel followed by reverse phase chromatography (C_{18}). Oligonucleotides 22, 23, 24 and 25 were kinased before annealing to their complementary oligonucleotides. The oligonucleotides were ligated pair E to pair F to pair G to form the gene coding for part of SCR3 and the whole of SCR4 which for convenience will be called the SCR4 gene in the subsequent text.

The DNA coding for SCR4 was initially amplified by PCR using two oligonucleotides (Table 1; oligos 17 and 18) complementary to the two strands of DNA. B th oligonucleotides c ntained 5' unmatched ends that contained 6bp of random sequence followed by the sequence of either *Spe*I (oligo 17) or *Hind* III (olig 18) restriction sites followed by 18 bp complementary to the gene. Following PCR a band of approximately 250 bp was visualised on horizontal agarose gel

electrophoresis, which was excised and purified on DEAE NA45 membranes. This DNA was used for a second PCR amplification using nested primers that had been moved inwards by four nucle tides at their 5' ends (Table 1; oligo 19, lig 20). These oligo's incorporated the *Spe*I and *Hind*III restriction sites but now only had 2 nucleotides beyond the end of each restriction site. Following PCR a band of approx. 250 bp was visualised on horizontal agarose gel electrophoresis. This band was excised and purified using the QIAEX agarose gel extraction kit.

The DNA for SCR 4 was blunt-end ligated to itself following kinasing. The multimers formed were visualized by horizontal agarose gel electrophoresis and the bands excised and purified using the QIAEX agarose gel extraction kit. The DNA was then cut with *Spe* I and *Hind* III and ligated into pDB1013-5-4 that had been cut with the same enzymes to produce pDB1018 (Fig.3). The vector was transformed into *E.coli* HB101 made competent with calcium chloride. Mini-plasmid preparations were made and plasmid DNA analysed by digestion with *Nde* I, *Hind* III, *Stu* I, *Spe* I and *Kpn* I. The plasmids with the correct restriction maps were analysed by DNA sequencing of both strands across the gene encoding SCR4. Two plasmids were selected for further study. pDB1018-1, which encoded MQ1-S253 (consensus SCR1 to 4) and pDB1018-6, which encoded the R235H mutant of MQ1-S253. The amino acid sequences of the two polypeptides encoded by pDB1018-1 and pDB1018-6 are shown in Table 2.

Taking the first residue as being the A of the ATG initiating codon, DNA sequencing revealed that residue 600 of pDB1018-6 had been altered from G -> A. This is a silent mutation and does not alter the amino acid at this position.

(b) **Expression of MQ1-S253 from pDB1018-1**

pDB1018-1, constructed as described in Example 11a, was transformed into calcium chloride competent *E.coli* BL21(DE3). Single colonies were inoculated into universals containing 10ml of NZCYM medium and 75 µg/ml ampicillin and allowed to grow overnight at 37°C, 230 r.p.m. 3ml of overnight culture were used to inoculate each of 8 x 2 litre conical flasks containing 500ml of NZCYM medium, 150 µg/ml ampicillin; cultures were grown at 37°C, 230 r.p.m. until A_{600} reached 0.5 absorbance units. The cultures were induced with 1mM IPTG and allowed to grow for a further 3 hours under the same conditions. The cultures were centrifuged (approx. 7000g/10 mins/4°C) and the supernatants discarded. The cell pellets were stored at -40°C.

(c) **Isolation of solubilised inclusion bodies**

The frozen cell pellets of *E.coli* BL21(DE3) (pDB1018-1) each equivalent to 1 litre of culture prepared as described in Example 11b were allowed to thaw at 0-4°C over 2 hours. The pellets were resuspended in 50mM Tris/50mM NaCl/1mM

EDTA/0.1mM PMSF pH 8.0; 30ml for each litre pellet. Each suspension was transferred to a 100ml glass beaker and sonicated (Heat systems - Ultrasonics W380; 70 Watts, 50 x 50% pulse, pulse time = 5 seconds). The sonicates were pooled and immediately centrifuged (6,000 g/ 4°C/10 mins) and the supernatant discarded. The pellet containing the inclusion bodies was resuspended in 20 mM Tris/8 M urea/50mM 2-mercaptoethanol/1mM EDTA/0.1mM PMSF pH 8.5 (400 ml), thoroughly mixed and left static at room temperature (approx. 23°C) for 1 hour.

(d) **Purification of MQ1-S253 from the solubilised inclusion body.**

30ml of S-Sepharose FF that had been washed with deionised water and suction dried was added to the inclusion body solution described in Example 11c, and vigorously shaken for 30 seconds. The S-Sepharose mixture was left static at room temperature (23°C) for 1.5 hours and then the supernatant was discarded. The remaining slurry was packed into a column (id, 4.1cm). The column was equilibrated using 20mM Tris/8M urea/50mM 2-mercaptoethanol/1mM EDTA/0.1mM PMSF pH 8.5 at 60 cmh^{-1}, 4°C. MQ1-S253 protein was eluted using the equilibration buffer containing 1M NaCl. Analysis by SDS PAGE/protein staining of the fractions collected during the chromatography indicated that virtually all the target protein had adsorbed to the column and had been dissociated by the 1M NaCl wash. The appropriate fraction was stored at -40°C.

(e) **Folding and formulation**

Based on a molar extinction coefficient of 25,000 and A_{280} values determined in 50mM formic acid, 60 mg of the S-Sepharose purified unfolded protein described in Example 11d was folded and formulated as follows :-

8.0ml of solution (equivalent to 60 mg protein) was diluted with 22ml cold 20mM Tris/8M urea/50mM 2-mercaptoethanol/1M NaCl/1mM EDTA/0.1mM PMSF pH8.5, to give 30ml of a 2.0 mg/ml solution. The 30ml was diluted rapidly with constant stirring into 930ml cold (0-4°C) freshly prepared 20mM ethanolamine. The solution was left static at 0-4°C for 1 hour. Reduced glutathione was added to 1mM (by addition of 9.6ml of 100mM stock) and then oxidised glutathione was added to 0.5mM (by addition of 9.6ml of 50mM stock). The solution was left static at 0-4°C for a further 48 hours and then ultrafiltered using a stirred cell (Amicon) and a YM10 membrane (Amicon, nominal 10,000 Da molecular weight cut-off) to approx. 29ml. The ultrafiltered retentate was buffer exchanged into 50 mM formic acid using Sephadex G25 (i.d., 26mm; h, 245mm Vt, 123ml) and a flow rate of 50 cmh^{-1} to a final volume of 40ml. Using a molar extinction coefficient of 25,000 for the protein 51mg of protein was recovered. The purified protein gave an IH_{50} value (see Method xvi) of approximately 2 µg/ml.

(f) **Further purification and f rmulation of SCR1 + 2 + 3 + 4.**

Folded SCR1 + 2 + 3 + 4 (nominal 25mg) in 50mM formic acid prepared essentially as described in Example 11e was lyophilised. The lyophilisate was resolubilised in 20mM ethanolamine (10ml) to give a cloudy solution. The 10ml were then added to 90ml 0.1M NaH_2PO4/1M $(NH4)_2SO4$ pH 7.0, thoroughly mixed, and then clarified by centrifugation (4000 rpm/20 min). The supernatant (100ml) was decanted and was chromatographed on Butyl Toyopearl (exactly as described for SCR1 + 2 + 3 in Example 3d). The peak A_{280} fractions, eluting at about 100% of the 1M NaCl-containing buffer, were pooled and buffer-exchanged using Sephadex G25 into 50mM formic acid. The Vo pool (29.5ml) was lyophilised in aliquots.

The purity of the protein was assessed by SDS PAGE followed by protein staining and by C8 reverse-phase HPLC; the protein was estimated to be >95% pure. One of the lyophilised aliquots was resolubilised to 4mg protein/ml in 0.1M Hepes/0.15M NaCl pH7.4. The product showed activity in the haemolytic assay (Method (xvi)); the IH50 was calculated to be 0.3 µg/ml.

Another of the lyophilised aliquots was examined to determine the disulphide bridge pattern using proteolytic digestion and peptide identification by amino acid sequencing. All eight correct (as predicted on the basis of a consensus SCR motif) disulphides were detected.

Example 12 Expression, isolation, folding and formulation of purified MQ1-S253 (R235H)

(a) **Expression of MQ1-S253 (R235H)**

pDB1018-6 (prepared as described in Example 11a) was transformed into calcium chloride competent *E. coli* BL21(DE3). Single colonies were inoculated into universals containing 10 mls of NCYZM medium and 50 µg/ml ampicillin and allowed to grow overnight at 37°C, 220 r.p.m. The overnight cultures (approx. 3ml) were used to inoculate each of 2 l conical flasks containing 500 ml of NCYZM medium, 150 µg/ml ampicillin; cultures were grown at 37°C, 220 r.p.m. until A_{600} was 0.5 absorbance units. Cultures were induced with 1 mM IPTG and allowed to grow a further 3 hours under the same conditions. The cultures were centrifuged (approx. 8000g/10 min/4°C) and the supernatants discarded. The cell pellets were stored at -40°C.

(b) **Isolation of solubilised inclusion bodies and purification of unfolded MQ1-S253 (R235H)**

Frozen cell pellet of *E. coli* BL21 DE3 (pDB1018-6) (2 litre culture) described in Example 12a was all wed to thaw at 4°C for 2 h and was then resuspended in 50 mM Tris/50 mM NaCl/1 mM EDTA/0.1 mM PMSF pH 8.0 (66

ml). The suspension was transferred to a 250 ml glass beaker and sonicated (Heat Systems - Ultrasonics W380; 70 Watts, 30 x 50% pulse time = 5 seconds). The sonicate was immediately centrifuged (6000g/4°C/10 min) and the supernatant was discarded. The pellet, containing the inclusion bodies, was resuspended by vigorous swirling in 20 mM Tris/8 M urea/50 mM 2-mercaptoethanol/1 mM EDTA/0.1 mM PMSF pH 8.5 (200 ml) and left static at room temperature (approx. 23°C) for 1.5h. Water-washed, suction-dried S-Sepharose (equivalent to approx. 25 ml packed bed volume) was added to the 200 ml solubilised inclusion body and the mixture swirled vigorously to wet the Sepharose beads thoroughly. The mixture was left static at room temperature for 1h. The supernatant (approximately 150 ml) was decanted and discarded. The slurry remaining was resuspended to a uniform suspension by swirling and then poured into a 32 mm (i.d.) glass jacket and allowed to settle. The gel bed was connected into a low pressure chromatography system and was equilibrated with 20 mM Tris/8 M urea/1mM EDTA/50mM 2-mercaptoethanol pH 8.5 at 4°C until the A_{280} baseline stabilised. The column was then developed with equilibration buffer containing 1M NaCl. All the chromatography was at approx. 1 ml min^{-1}. Analysis by SDS PAGE/protein staining of the fractions collected during the chromatography indicated that most of the MQ1-S253 (R235H) polypeptide had adsorbed to the column and had been dissociated by the 1M NaCl - containing buffer wash and that the purity of the material was about 90%.

A sample of the pool was buffer-exchanged into 50mM formic acid using Sephadex G25 column to allow some assays to be carried out.

Amino acid analysis of the pool of the MQ1-S253 (R235H) - containing fractions gave a total protein content of about 120 mg.

(c) **Folding and formulation of SCR 1 + 2 + 3 (R235H)**

Based on A_{280} values and a molar extinction coefficient of 25,000 for the protein in 50mM formic acid, 20 mg of the S-Sepharose-purified unfolded protein described in Example 12b was folded and formulated as follows.

5.2 ml protein solution (equivalent to 20 mg) was diluted with 4.8 ml cold 20 mM Tris/8M urea/50 mM 2-mercaptoethamol/1M NaCl pH8.5, to yield 10 ml of a 2.0 mg/ml solution.

The 10 ml was diluted rapidly with constant stirring into 310 ml freshly prepared, cold (approx. 0-4°C) 20 mM ethanolamine. The solution was left static at 0-4°C for 1 h. Then reduced glutathione was added to 1 mM (by addition of 2.56 ml 125 mM GSH). Then oxidised glutathione was added to 0.5 mM (by addition of 3.2 ml 50 mM GSSG). The solution was left static, in the cold room (~ 2-3°C), for a further 48 h. The solution was then ultrafiltered using a stirred cell and a YM10 (nominal 10,000 molecular weight out-off) membrane to approximately 2 ml. The

solution was clear. The ultrafiltration cell was washed with approximately 2 ml 20 mM ethanolamine and the wash and the ultrafiltered retentate were pooled to give a final volume of 3.7 ml.

2.2 ml of this solution was buffer-exchanged into 3.2 ml 50 mM formic acid using Sephadex G25 (PD10). The buffer-exchanged material was regarded as the product, and it was stored at -40°C. Analysis of an aliquot of the product showed it contained 1.6 mg protein/ml, that by SDS PAGE under non-reducing conditions a single major band of M_r~28,000 was present and that N-terminal sequencing of the band (MQXNAPE) was consistent with the expected sequence. In addition the preparation gave an IH_{50} value (see Method (xvi)) of approximately 1 µg/ml.

IN THE FIGURES

Fig. 1 Plasmid pBROC413. *bla* indicates the ampicillin resistance gene, ø10 the T7 RNA polymerase promoter and rbs the ribosome binding site. Arrows for ø10 and *bla* give the direction of transcription. The polylinker site has been indicated. The plasmid is not drawn to scale and the size is approximate.

Figure 2 illustrates the construction from pDB1010-D11 and pBROC435 of plasmid pDB1013-5-4 coding for SCR 1 + 2 + 3. Plasmid sizes are approximate and are not drawn to scale.

Figure 3 illustrates the construction from pDB1013-5-4 of pDB1018 coding for SCR 1+2+3+4. Plasmid sizes are approximate and are not drawn to scale.

REFERENCES USED IN EXAMPLES OR GENERAL METHODS

1. Chen G-F. and Inouye M. (1990). Suppression of the negative effect of minor arginine codons on gene expression; preferential usage of minor codons within the first 25 codons of the *E.coli* genes. *Nuc.Acids.Res. 18 (6)*: 1465-1473.

2. Dacie, J.V. & Lewis S.M., (1975) *Practical haematology* Fifth Edition, Ed. Churchill Livingstone, Edinburgh & New York pp 3 - 4

3. Devereux, Haeberli and Smithies (1984). A comprehensive set of sequence analysis programmes for the vax. *Nuc.Acids.Res. 12(1)*: 387 - 395

4. Hourcade D., Miesner D. R., Atkinson J. P. & Holers V. M. (1988). Identification of an alternative polyadenylation site in the human C3b/C4b receptor (complement receptor type 1) transcriptional unit and prediction of a secreted form of complement receptor type 1. *J. Exp. Med.* **168** 1255 - 1270

5. Low B. (1968). *Proc.Natl.Acad.Sci.USA 60*: 160

6. Sambrook J., Fritsch E.F. and Maniatis J. (1989). Molecular Cloning: A Laboratory Manual 2nd Edition. Cold Spring Harbour Laboratory Press

7. Studier F.W. and Moffat B.A. (1986). *J.Mol.Biol 189*: 113

8. Tabor S. (1990). Expression using the T7 RNA polymerase/promoter system. In Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, .D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl, eds) pp. 16.2.1 - 16.2.11 Greene Publishing and Wiley-Interscience, New York

TABLE 1

OLIGO 1 = SEQ ID NO: 1
OLIGO 2 = SEQ ID NO: 2
OLIGO 3 = SEQ ID NO: 3
OLIGO 4 = SEQ ID NO: 4
OLIGO 5 = SEQ ID NO: 5
OLIGO 6 = SEQ ID NO: 6
OLIGO 7 = SEQ ID NO: 7
OLIGO 8 = SEQ ID NO: 8
OLIGO 9 = SEQ ID NO: 9
OLIGO 10 = SEQ ID NO: 10
OLIGO 11 = SEQ ID NO: 11
OLIGO 12 = SEQ ID NO: 12
OLIGO 13 = SEQ ID NO: 13
OLIGO 14 = SEQ ID NO: 14
OLIGO 15 = SEQ ID NO: 15
OLIGO 16 = SEQ ID NO: 16
OLIGO 17 = SEQ ID NO: 17
OLIGO 18 = SEQ ID NO: 18
OLIGO 19 = SEQ ID NO: 19
OLIGO 20 = SEQ ID NO: 20
OLIGO 21 = SEQ ID NO: 21
OLIGO 22 = SEQ ID NO: 22
OLIGO 23 = SEQ ID NO: 23
OLIGO 24 = SEQ ID NO: 24
OLIGO 25 = SEQ ID NO: 25
OLIGO 26 = SEQ ID NO: 26

TABLE 2 Amino acid sequences of proteins, deduced from the cDNA constructs.

The full deduced sequence of the proteins of the Examples are given as follows:

MQ1->K196 of CR-1 is given in SEQ ID NO: 27

MR122->K196 of CR-1 is given in SEQ ID NO: 28

MQ1-S253 of CR-1 is given in SEQ ID NO: 29

The R235H mutant of MQ1-S253 of CR-1 is given in SEQ ID NO: 30

MR122-S253 of CR-1 is given in SEQ ID NO: 31.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: SmithKline Beecham p.l.c. -, - -

(ii) TITLE OF INVENTION: Novel Compounds

(iii) NUMBER OF SEQUENCES: 31

(iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: SmithKline Beecham Corporate Patents
- (B) STREET: Great Burgh, Yew Tree Bottom Road
- (C) CITY: Epsom
- (D) STATE: Surrey
- (E) COUNTRY: England
- (F) ZIP: KT18 5XQ

(v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER:
- (B) FILING DATE:
- (C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: Valentine, Jill B
- (B) REGISTRATION NUMBER: G.A. 26758
- (C) REFERENCE/DOCKET NUMBER: P30423

(ix) TELECOMMUNICATION INFORMATION:
- (A) TELEPHONE: 0737364158

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 87 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TATGCAGTGC AACGCTCCGG AATGGCTGCC GTTCGCGCGC CCGACCAACC TGACTGATGA 60

ATTTGAGTTC CCGATCGGTA CCTACCT 87

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 93 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CGTAGTTCAG GTAGGTACCG ATCGGGAACT CAAATTCATC AGTCAGGTTG GTCGGGCGCG 60

CGAACGGCAG CCATTCCGGA GCGTTGCACT GCA 93

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 101 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAACTACGAA TGCCGCCCGG GTTATAGCGG CCGCCCGTTT TCTATCATCT GCCTGAAAAA 60

CTCTGTCTGG ACTGGTGCTA AGGACCGTTG CCGACGTAAA T 101

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 101 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ACGACAAGAT TTACGTCGGC AACGGTCCTT AGCACCAGTC CAGACAGAGT TTTTCAGGCA 60

GATGATAGAA AACGGGCGGC CGCTATAACC CGGGCGGCAT T 101

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 101 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTTGTCGTAA TCCGCCAGAT CCGGTTAACG GCATGGTGCA TGTGATCAAA GGCATCCAGT 60

TCGGTTCCCA AATTAAATAT TCTTGTACTA AAGGTTACCG T 101

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 101 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCAATCAGAC GGTAACCTTT AGTACAAGAA TATTTAATTT GGGAACCGAA CTGGATGCCT 60

TTGATCACAT GCACCATGCC GTTAACCGGA TCTGGCGGAT T 101

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 94 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTGATTGGTT CCTCCAGCGC TACATGCATC ATCTCTGGTG ATACTGTCAT TTGGGATAAT 60

GAAACACCGA TTTGTGACCG AATTCAGTAA TAAA 94

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 90 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGCTTTTATT ACTGAATTCG GTCACAAATC GGTGTTTCAT TATCCCAAAT GACAGTATCA 60

CCAGAGATGA TGCATGTAGC GCTGGAGGAA 90

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 72 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TATGCGAATT CCGTGTGGTC TGCCGCCGAC CATCACCAAC GGTGATTTCA TCTCTACCAA 60

TCGCGAGAAT TT 72

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 78 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CATAGTGAAA ATTCTCGCGA TTGGTAGAGA TGAAATCACC GTTGGTGATG GTCGGCGGCA 60

GACCACACGG AATTCGCA 78

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 85 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TCACTATGGT TCTGTGGTGA CCTACCGCTG CAATCCGGGT AGCGGTGGTC GTAAGGTGTT 60

TGAGCTCGTG GGTGAGCCGT CCATC 85

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 85 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GTGCAGTAGA TGGACGGCTC ACCCACGAGC TCAAACACCT TACGACCACC GCTACCCGGA 60

TTGCAGCGGT AGGTCACCAC AGAAC 85

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 79 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TACTGCACTA GTAATGACGA TCAAGTGGGC ATCTGGAGCG GCCCGGCACC GCAGTGCATC 60

ATCCCGAACA AATAATAAA 79

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 75 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AGCTTTTATT ATTTGTTCGG GATGATGCAC TGCGGTGCCG GGCCGCTCCA GATGCCCACT 60

TGATCGTCAT TACTA 75

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GAGACTCATA TGCAGTGCAA CGCTCCGGAA 30

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GTCAGCAAGC TTTTATTACT GAATTCGGTC 30

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ATCGTAACTA GTAACGACGA TCAAGTGGGC 30

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ATGACTAAGC TTTTATTATG AGCAGCTCGG 30

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TAACTAGTAA CGACGATCAA GTGGGCATCT GG 32

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CTAAGCTTTT ATTATGAGCA GCTCGGGAGT TCC 33

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 81 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CTAGTAACGA CGATCAAGTG GGCATCTGGA GCGGCCCGGC ACCGCAGTGC ATCATCCCGA 60

ACAAATGCAC GCCGCCAAAT G 81

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 85 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GTTCTCCACA TTTGGCGGCG TGCATTTGTT CGGGATGATG CACTGCGGTG CCGGGCCGCT 60

CCAGATGCCC ACTTGATCGT CGTTA 85

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 90 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TGGAGAACGG TATCCTGGTA TCTGACAACC GTTCTCTGTT CTCTTTAAAC GAAGTTGTAG 60

AGTTTCGTTG TCAGCCGGGC TTTGTTATGA 90

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 90 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CGGACCTTTC ATAACAAAGC CCGGCTGACA ACGAAACTCT ACAACTTCGT TTAAAGAGAA 60

CAGAGAACGG TTGTCAGATA CCAGGATACC 90

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 72 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

AAGGTCCGCG CCGTGTGAAG TGCCAGGCCT TGAACAAATG GGAGCCGGAA CTCCCGAGCT 60

GCTCATAATA AA 72

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AGCTTTTATT ATGAGCAGCT CGGGAGTTCC GGCTCCCATT TGTTCAAGGC CTGGCACTTC 60

ACACGGCG 68

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 197 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Met Gln Cys Asn Ala Pro Glu Trp Leu Pro Phe Ala Arg Pro Thr Asn
1 5 10 15

Leu Thr Asp Glu Phe Glu Phe Pro Ile Gly Thr Tyr Leu Asn Tyr Glu
20 25 30

Cys Arg Pro Gly Tyr Ser Gly Arg Pro Phe Ser Ile Ile Cys Leu Lys
35 40 45

Asn Ser Val Trp Thr Gly Ala Lys Asp Arg Cys Arg Arg Lys Ser Cys
50 55 60

Arg Asn Pro Pro Asp Pro Val Asn Gly Met Val His Val Ile Lys Gly
65 70 75 80

Ile Gln Phe Gly Ser Gln Ile Lys Tyr Ser Cys Thr Lys Gly Tyr Arg
85 90 95

Leu Ile Gly Ser Ser Ser Ala Thr Cys Ile Ile Ser Gly Asp Thr Val
100 105 110

Ile Trp Asp Asn Glu Thr Pro Ile Cys Asp Arg Ile Pro Cys Gly Leu
115 120 125

Pro Pro Thr Ile Thr Asn Gly Asp Phe Ile Ser Thr Asn Arg Glu Asn
130 135 140

Phe His Tyr Gly Ser Val Val Thr Tyr Arg Cys Asn Pro Gly Ser Gly
145 150 155 160

Gly Arg Lys Val Phe Glu Leu Val Gly Glu Pro Ser Ile Tyr Cys Thr
165 170 175

Ser Asn Asp Asp Gln Val Gly Ile Trp Ser Gly Pro Ala Pro Gln Cys
180 185 190

Ile Ile Pro Asn Lys
195

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 76 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Met Arg Ile Pro Cys Gly Leu Pro Pro Thr Ile Thr Asn Gly Asp Phe
1 5 10 15

Ile Ser Thr Asn Arg Glu Asn Phe His Tyr Gly Ser Val Val Thr Tyr
20 25 30

Arg Cys Asn Pro Gly Ser Gly Gly Arg Lys Val Phe Glu Leu Val Gly
35 40 45

Glu Pro Ser Ile Tyr Cys Thr Ser Asn Asp Asp Gln Val Gly Ile Trp
50 55 60

Ser Gly Pro Ala Pro Gln Cys Ile Ile Pro Asn Lys
65 70 75

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 254 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Met Gln Cys Asn Ala Pro Glu Trp Leu Pro Phe Ala Arg Pro Thr Asn
1 5 10 15

Leu Thr Asp Glu Phe Glu Phe Pro Ile Gly Thr Tyr Leu Asn Tyr Glu
20 25 30

Cys Arg Pro Gly Tyr Ser Gly Arg Pro Phe Ser Ile Ile Cys Leu Lys
35 40 45

Asn Ser Val Trp Thr Gly Ala Lys Asp Arg Cys Arg Arg Lys Ser Cys

50 55 60

Arg Asn Pro Pro Asp Pro Val Asn Gly Met Val His Val Ile Lys Gly
65 70 75 80

Ile Gln Phe Gly Ser Gln Ile Lys Tyr Ser Cys Thr Lys Gly Tyr Arg
85 90 95

Leu Ile Gly Ser Ser Ser Ala Thr Cys Ile Ile Ser Gly Asp Thr Val
100 105 110

Ile Trp Asp Asn Glu Thr Pro Ile Cys Asp Arg Ile Pro Cys Gly Leu
115 120 125

Pro Pro Thr Ile Thr Asn Gly Asp Phe Ile Ser Thr Asn Arg Glu Asn
130 135 140

Phe His Tyr Gly Ser Val Val Thr Tyr Arg Cys Asn Pro Gly Ser Gly
145 150 155 160

Gly Arg Lys Val Phe Glu Leu Val Gly Glu Pro Ser Ile Tyr Cys Thr
165 170 175

Ser Asn Asp Asp Gln Val Gly Ile Trp Ser Gly Pro Ala Pro Gln Cys
180 185 190

Ile Ile Pro Asn Lys Cys Thr Pro Pro Asn Val Glu Asn Gly Ile Leu
195 200 205

Val Ser Asp Asn Arg Ser Leu Phe Ser Leu Asn Glu Val Val Glu Phe
210 215 220

Arg Cys Gln Pro Gly Phe Val Met Lys Gly Pro Arg Arg Val Lys Cys
225 230 235 240

Gln Ala Leu Asn Lys Trp Glu Pro Glu Leu Pro Ser Cys Ser
245 250

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 254 amino acids

(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

```
Met Gln Cys Asn Ala Pro Glu Trp Leu Pro Phe Ala Arg Pro Thr Asn
1               5                   10                  15

Leu Thr Asp Glu Phe Glu Phe Pro Ile Gly Thr Tyr Leu Asn Tyr Glu
            20                  25                  30

Cys Arg Pro Gly Tyr Ser Gly Arg Pro Phe Ser Ile Ile Cys Leu Lys
        35                  40                  45

Asn Ser Val Trp Thr Gly Ala Lys Asp Arg Cys Arg Arg Lys Ser Cys
    50                  55                  60

Arg Asn Pro Pro Asp Pro Val Asn Gly Met Val His Val Ile Lys Gly
65                  70                  75                  80

Ile Gln Phe Gly Ser Gln Ile Lys Tyr Ser Cys Thr Lys Gly Tyr Arg
                85                  90                  95

Leu Ile Gly Ser Ser Ser Ala Thr Cys Ile Ile Ser Gly Asp Thr Val
            100                 105                 110

Ile Trp Asp Asn Glu Thr Pro Ile Cys Asp Arg Ile Pro Cys Gly Leu
        115                 120                 125

Pro Pro Thr Ile Thr Asn Gly Asp Phe Ile Ser Thr Asn Arg Glu Asn
    130                 135                 140

Phe His Tyr Gly Ser Val Val Thr Tyr Arg Cys Asn Pro Gly Ser Gly
145                 150                 155                 160

Gly Arg Lys Val Phe Glu Leu Val Gly Glu Pro Ser Ile Tyr Cys Thr
                165                 170                 175
```

```
Ser Asn Asp Asp Gln Val Gly Ile Trp Ser Gly Pro Ala Pro Gln Cys
            180                 185                 190

Ile Ile Pro Asn Lys Cys Thr Pro Pro Asn Val Glu Asn Gly Ile Leu
        195                 200                 205

Val Ser Asp Asn Arg Ser Leu Phe Ser Leu Asn Glu Val Val Glu Phe
    210                 215                 220

Arg Cys Gln Pro Gly Phe Val Met Lys Gly Pro His Arg Val Lys Cys
225                 230                 235                 240

Gln Ala Leu Asn Lys Trp Glu Pro Glu Leu Pro Ser Cys Ser
                245                 250
```

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 133 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```
Met Arg Ile Pro Cys Gly Leu Pro Pro Thr Ile Thr Asn Gly Asp Phe
1               5                   10                  15

Ile Ser Thr Asn Arg Glu Asn Phe His Tyr Gly Ser Val Val Thr Tyr
            20                  25                  30

Arg Cys Asn Pro Gly Ser Gly Gly Arg Lys Val Phe Glu Leu Val Gly
        35                  40                  45

Glu Pro Ser Ile Tyr Cys Thr Ser Asn Asp Asp Gln Val Gly Ile Trp
    50                  55                  60

Ser Gly Pro Ala Pro Gln Cys Ile Ile Pro Asn Lys Cys Thr Pro Pro
65                  70                  75                  80
```

```
Asn Val Glu Asn Gly Ile Leu Val Ser Asp Asn Arg Ser Leu Phe Ser
                85                  90                  95

Leu Asn Glu Val Val Glu Phe Arg Cys Gln Pro Gly Phe Val Met Lys
            100                 105                 110

Gly Pro Arg Arg Val Lys Cys Gln Ala Leu Asn Lys Trp Glu Pro Glu
        115                 120                 125

Leu Pro Ser Cys Ser
    130
```

Claims

1. A soluble polypeptide comprising, in sequence, one to four short consensus repeats (SCR) selected from SCR 1, 2, 3 and 4 of long homologous repeat A (LHR-A) as the only structurally and functionally intact SCR domains of CR1 and including at least SCR3.

2. A polypeptide according to claim 1 which comprises, in sequence, SCR 1, 2 and 3 of LHR-A as the only structurally and functionally intact SCR domains of CR1.

3. A polypeptide according to claim 2 of formula (I):

$$NH_2\text{-}V^1\text{-}SCR1\text{-}W^1\text{-}SCR2\text{-}X^1\text{-}SCR3\text{-}Y^1\text{-}OH \qquad (I)$$

in which SCR1 represents residues 2-58 of mature CR1, SCR2 represents residues 63-120 of mature CR1, SCR3 represents residues 125-191 of mature CR1, and V^1, W^1, X^1 and Y^1 represent bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

4. A polypeptide according to claim 3 in which W^1, X^1 and Y^1 represent residues 59-62, 121-124 and 192-196, respectively, of mature CR1 and V^1 represents residue 1 of mature CR1 optionally linked via its N-terminus to methionine.

5. A polypeptide according to claim 1 which comprises, in sequence, SCR 1, 2, 3 and 4 of LHR-A as the only structurally and functionally intact SCR domains of CR1.

6. A polypeptide according to claim 5 of formula (II):

$$NH_2\text{-}V^2\text{-}SCR1\text{-}W^2\text{-}SCR2\text{-}X^2\text{-}SCR3\text{-}Y^2\text{-}SCR4\text{-}Z^2OH \qquad (II)$$

in which SCR1, SCR2 and SCR3 are as defined in claim 3, SCR4 represents residues 197-252 of mature CR1 and V^2, W^2, X^2, Y^2 and Z^2 represents bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

7. A polypeptide according to claim 6 in which W^2, X^2, Y^2 and Z^2 represent residues 59-62, 121-124, 192-196, and residues 253 respectively, of mature CR1 and V^2 represents residue 1 of mature CR1 optionally linked via its N-terminus to methionine.

8. A polypeptide according to claim 7 in which arginine 235 is replaced by histidine.

9. A polypeptide according to claim 1 of formula (III):

$$NH_2\text{-}X^3\text{-}SCR3\text{-}Y^3\text{-}OH \qquad (III)$$

in which SCR3 is as defined in claim 3 and X^3 and Y^3 represent bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

10. A polypeptide according to claim 9 in which X^3 represents amino acids 122-124 of mature CR1 optionally linked to methionine at its N-terminus and Y^4 represents amino acids 192-196 of mature CR1.

11. A polypeptide according to claim 1 of formula (IV):

$$NH_2\text{-}X^4\text{-}SCR3\text{-}Y^4\text{-}SCR4\text{-}Z^4\text{-}OH \qquad \text{(IV)}$$

in which SCR3 and SCR4 are as defined in claim 6 and X^4, Y^4 and Z^4 represent bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

12. A polypeptide according to claim 11 in which X^4 represents amino acids 122-124 of mature CR1 optionally linked to methionine at its N-terminus and Y^4 and Z^4 represent amino acids 192-196 and 253 respectively of mature CR1.

13. A polypeptide having the amino acid sequence given in SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30 or SEQ ID NO:31.

14. A DNA polymer comprising a nucleotide sequence that encodes the polypeptide according to claim 1.

15. A replicable expression vector capable, in a host cell, of expressing the DNA polymer of claim14.

16. A host cell transformed with the replicable expression vector of claim 15.

17. A process for preparing a CR1 polypeptide according to claim 1 which process comprises expressing DNA encoding said polypeptide in a recombinant host cell and recovering the product.

18. A pharmaceutical composition comprising a therapeutically effective amount of a polypeptide according to claim 1 and a pharmaceutically acceptable carrier or excipient.

19. A polypeptide according to claim 1 for use as an active therapeutic substance.

20. A method of treating a disease or disorder associated with inflammation or inappropriate complement activation comprising administering to a subject in need of such treatment a therapeutically effective amount of a polypeptide according to claim 1.

21. The use of a polypeptide of claim 1 in the manufacture of a medicament for the treatment of a disease or disorder associated with inflammation or inappropriate complement activation.

Figure 1

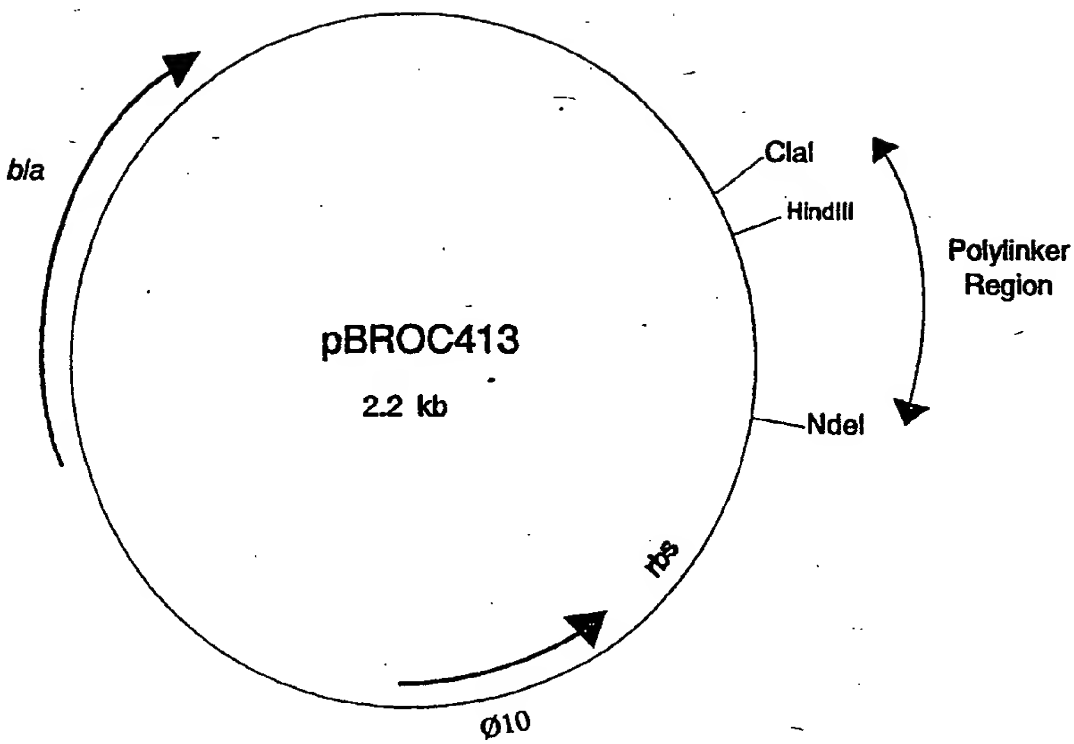

bla
ClaI
HindIII
Polylinker Region
pBROC413
2.2 kb
NdeI
rbs
Ø10

Figure 2

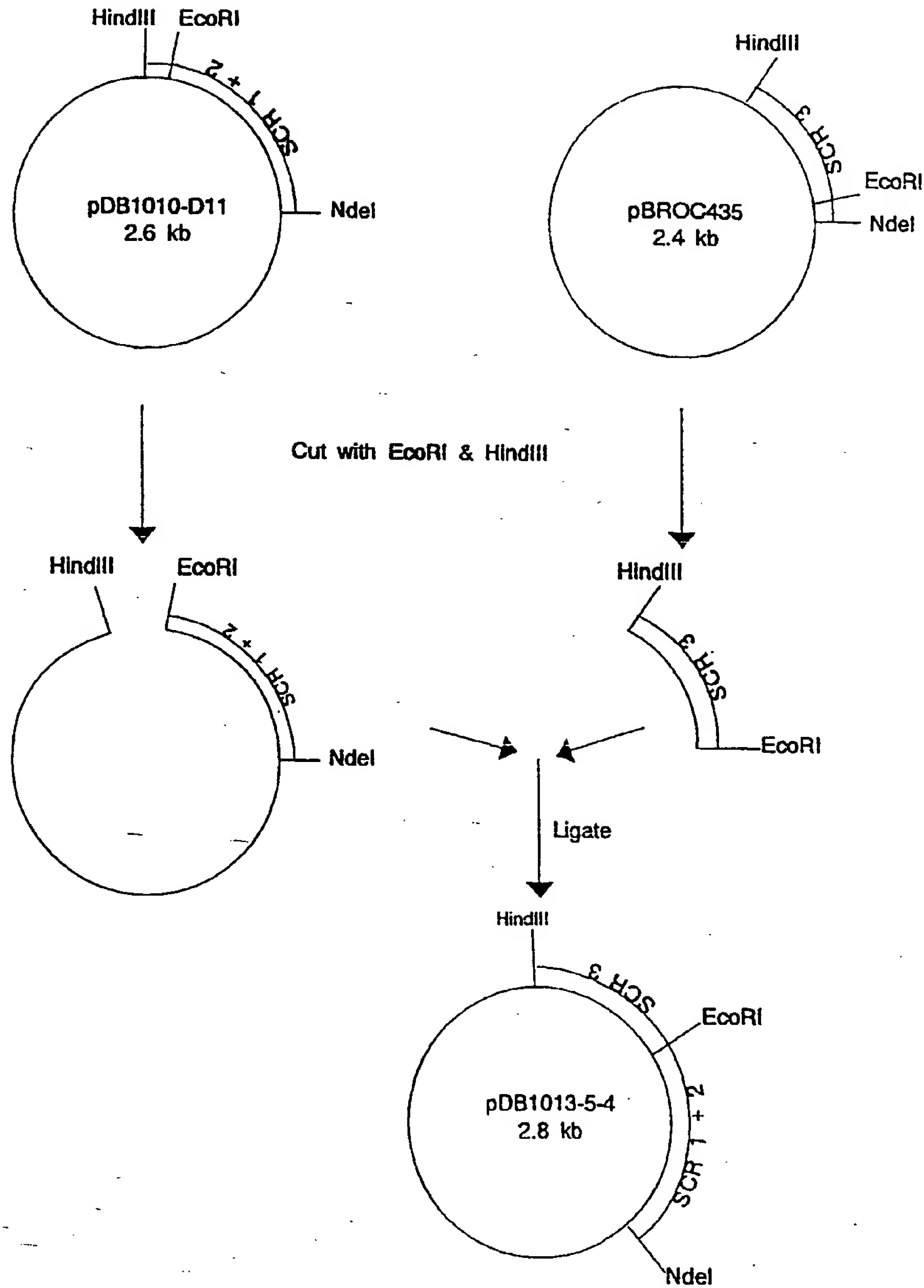
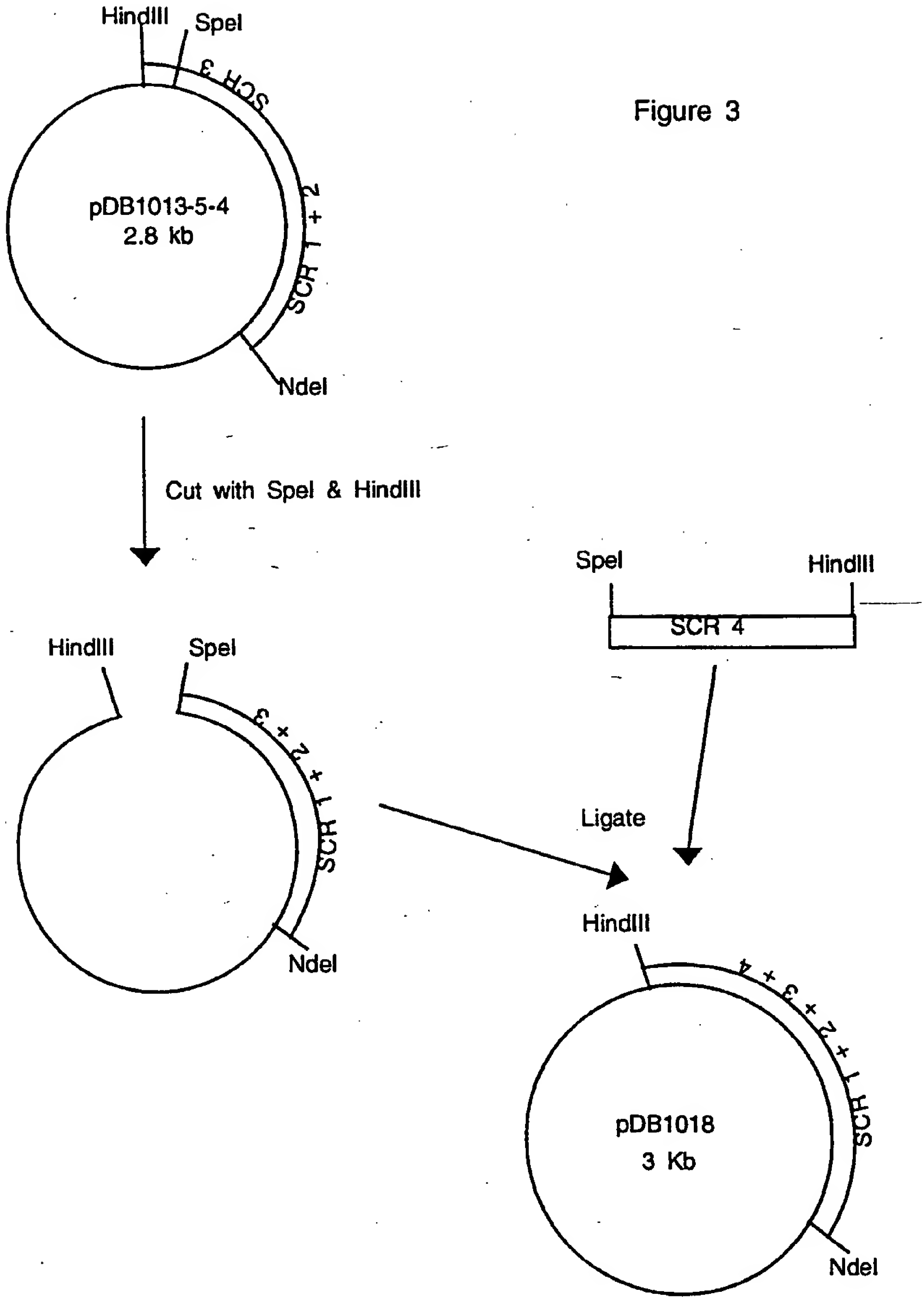

HindIII
EcoRI
SCR 1 + 2
NdeI
pDB1010-D11
2.6 kb
HindIII
SCR 3
EcoRI
NdeI
pBROC435
2.4 kb
Cut with EcoRI & HindIII
HindIII
EcoRI
SCR 1 + 2
NdeI
HindIII
SCR 3
EcoRI
Ligate
HindIII
SCR 3
EcoRI
SCR 1 + 2
pDB1013-5-4
2.8 kb
NdeI

Figure 3


HindIII
SpeI
SCR 3
SCR 1 + 2
pDB1013-5-4
2.8 kb
NdeI
Cut with SpeI & HindIII
HindIII
SpeI
SCR 1 + 2 + 3
NdeI
SpeI
HindIII
SCR 4
Ligate
HindIII
SCR 1 + 2 + 3 + 4
pDB1018
3 Kb
NdeI

INTERNATIONAL SEARCH REPORT

International Application No PCT/GB 93/01282

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[6]

According to International Patent Classification (IPC) or to both National Classification and IPC

Int.Cl. 5 C12N15/12; A61K37/02; C07K13/00; C12N1/21
//(C12N1/21,C12R1:19)

II. FIELDS SEARCHED

Minimum Documentation Searched[7]

Classification System	Classification Symbols
Int.Cl. 5	C07K ; C12N ; A61K

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched[8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT[9]

Category°	Citation of Document,[11] with indication, where appropriate, of the relevant passages[12]	Relevant to Claim No.[13]
A	WO,A,8 909 220 (THE JOHNS HOPKINS UNIVERSITY ET AL.) 5 October 1989 cited in the application see pages 32-38, chapters 5.5 and 5.6 see example 11.2.4.3 see figure 20 ---	1-21
A	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA vol. 88, no. 10, 15 May 1991, WASHINGTON US pages 4353 - 4357 KRYCH, M. ET AL. 'Sites within the complement C3b/C4b receptor important for the specificity of ligand binding' cited in the application see the whole document ---	1-21

-/--

° Special categories of cited documents :[10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

2

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
13 OCTOBER 1993	22.10.93
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	ANDRES S.M.

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)		
Category °	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No.
A	BIOLOGICAL ABSTRACTS vol. 93 , 1992, Philadelphia, PA, US; abstract no. 53233, KALLI, K. ET AL. 'Mapping of the C3b-binding site of CR1 and construction of a (CR1)2-F(ab)'12 chimeric complement inhibitor' see abstract & J.EXP.MED., 174(6), pp.1451-60; 1991	1-21

INTERNATIONAL SEARCH REPORT

International application No.
PCT/GB93/01282

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark:Although claim 20 is directed to a method of treatment of the human/animal body the search has been carried out and based on the alleged effects of the compound/composition

2. [ ] Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. [ ] As all searchable claims could be searches without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO. GB 9301282
SA 75972

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report.
The members are as contained in the European Patent Office EDP file on 13/10/93
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-8909220	05-10-89	AU-A- 3539489	16-10-89
		EP-A- 0411031	06-02-91
		JP-T- 4501502	19-03-92
		US-A- 5212071	18-05-93

EPO FORM P0479

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

-continued

<220>
<223> Description of Artificial Sequence: commercial polynucleotide linker

<400> 7
gatgaactag tctcgagag 19

<210> 8
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: commercial polynucleotide linker

<400> 8
ctatcgagac tagttcatc 19

<210> 9
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: commercial phosphorylated mutagenic oligonucleotide

<400> 9
ggggttttac agtgaccatc gcgaacagaa agttcacagc 40

<210> 10
<211> 13
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic oligonucleotide linker

<400> 10
cgcttaagct cga 13

<210> 11
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic oligonucleotide linker

<400> 11
gatctcgagc ttaagcg 17

What is claimed is:

1. A composition comprising a soluble complement regulatory protein moiety which has a short consensus repeat structural motif and which binds a complement component, and a carbohydrate moiety which binds a cellular adhesion molecule selected from the group consisting of P-selectin, E-selectin, and L-selectin.

2. The composition of claim 1 in which said carbohydrate moiety includes carbohydrate structures selected from the group consisting of sLe^x, Le^x, sLe^a and Le^a.

3. The composition according to claim 1 which inhibits a primary event in an inflammatory response.

4. The composition according to claim 1 in which the carbohydrate moiety binds to activated endothelial cells.

5. The composition according to claim 1 which is capable of inhibiting at least one activity associated with complement.

6. The composition according to claim 1 wherein said carbohydrate moiety has a binding affinity of at least about 10^6 M^{-1}.

7. The composition according to claim 1 in which said carbohydrate moiety comprises at least one N-linked carbohydrate, which carbohydrate is fucosylated.

8. A composition comprising a soluble complement receptor protein or functional derivative or analog thereof capable of binding a complement component, and a carbohydrate moiety which is a ligand for a cellular adhesion molecule selected from the group consisting of P-selectin, E-selectin, and L-selectin.

9. A composition comprising soluble complement receptor type 1 or a functional derivative or analog thereof capable of binding a complement component, and a carbohydrate moiety which is a ligand for a cellular adhesion molecule selected from the group consisting of P-selectin, E-selectin, and L-selectin.

10. The composition of claim 9 in which said soluble complement receptor type 1, or derivative or analog thereof, comprises at least LHR B, LHR C, and LHR D, up to and including the first alanine residue of the transmembrane region of complement receptor type 1.

11. A pharmaceutical composition comprising a soluble complement regulatory protein moiety which has a short consensus repeat structural motif and which binds a complement component, and a carbohydrate moiety which binds a cellular adhesion molecule selected from the group consisting of P-selectin, E-selectin, and L-selectin, dispersed in a pharmaceutically acceptable carrier.

12. The pharmaceutical composition of claim 11 wherein said carbohydrate moiety includes a carbohydrate structure selected from the group consisting of sLe^x, Le^x, sLe^a and Le^a.

13. A soluble complement inhibitory protein comprising a selectin ligand.

14. The complement inhibitory protein of claim 13 in which the selectin ligand is a Lewis X antigen.

15. The complement inhibitory protein of claim 14 in which the Lewis X antigen is a sialyl Lewis X antigen.

16. The complement inhibitory protein of claim 14 which is selected from the group consisting essentially of soluble complement receptor 1, a fragment, a derivative, and an analog thereof.

17. The complement inhibitory protein of claim 15 which is soluble complement receptor 1 or a fragment, derivative or analog thereof.

18. A method for producing a composition comprising a complement moiety and a carbohydrate moiety, which comprises:

(a) expressing a complement related protein in a cell which glycosylates the complement related protein with a Lewis X antigen; and

(b) recovering the complement inhibitory protein expressed in step (a).

19. The method for producing a composition comprising a complement moiety and a carbohydrate moiety of claim 18 wherein said complement related protein is glycosylated with a sialylated Lewis X antigen.

20. The method of claim 18 in which the complement related protein is complement receptor 1, or an analog or derivative thereof.

21. The method of claim 19 in which the complement inhibitory protein is complement receptor 1, or an analog or derivative thereof.

22. A method for producing a complement inhibitory protein comprising a selectin ligand, which comprises linking a complement inhibitory protein with a selectin ligand.

23. The method of claim 22 in which the complement inhibitory protein is complement receptor 1.

24. The method for producing a complement inhibitory protein of claim 22 wherein said selectin ligand is a Lewis X antigen.

25. The method of claim 24 wherein said Lewis X antigen is sialylated.

26. The method of claim 22 further comprising the step of linking a carrier protein to said carbohydrate moiety.

27. The method of claim 26 in which the carrier protein is a short consensus repeat.

28. A method for treating a subject with a disease or disorder involving complement activity comprising administering to a subject in need of such treatment a complement inhibiting amount of a composition comprising a soluble complement regulatory protein moiety which has a short consensus repeat structural motif and which binds a complement component, and at least one carbohydrate moiety which is a ligand for a cellular adhesion molecule selected from the group consisting of P-selectin, E-selectin and L-selectin.

29. The method for treating a subject of claim 28 wherein said complement inhibitory protein comprises a Lewis X antigen.

30. The method for treating a subject of claim 29 complement inhibitory protein comprises a sialyl Lewis X antigen.

31. A method for diagnosing an inflammatory condition comprising the steps of:

1) immobilizing a composition comprising a complement moiety and a carbohydrate moiety, and

2) incubating said immobilized composition with a sample suspected of containing a ligand for the composition of step (1) so immobilized, and

3) detecting the amount of binding that has occurred.

32. A method for producing a composition that inhibits an activity associated with complement, which comprises:

(a) expressing a soluble protein moiety having a short consensus repeat motif in a cell which glycosylates said protein moiety with a carbohydrate that is a selectin ligand; and

(b) recovering a complement inhibitory composition from said cell.

33. The method according to claim 32, wherein said carbohydrate is a sialylated Lewis X antigen.

34. The method according to claim 32, wherein said protein moiety is a soluble complement receptor type 1, or an analog or derivative thereof which binds a complement component.

35. A method for treating a subject with an inflammatory condition comprising administering to a subject in need of such treatment an inflammation inhibiting amount of a composition comprising a soluble complement regulatory protein moiety which has a short consensus repeat structural motif and which binds a complement component, linked to a carbohydrate moiety which binds a cellular adhesion molecule selected from the group consisting of P-selectin, E-selectin and L-selectin.

36. A method for diagnosing an inflammatory condition comprising:

(a) contacting a composition according to claim 1 with a fluid or tissue sample from a subject suspected of having such condition, and

(b) detecting binding of the composition with cellular adhesion molecule components in the sample.

37. A method for diagnosing complement activation comprising:

(a) contacting a composition according to claim 1 with a fluid or tissue sample from a subject suspected of having activated complement, and

(b) detecting binding of the composition with complement components in the sample.

* * * * *

612,314

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12N 15/12, A61K 37/02, C07K 13/00, C12N 1/21 // (C12N 1/21, C12R 1:19)

A1

(11) International Publication Number: WO 94/00571

(43) International Publication Date: 6 January 1994 (06.01.94)

(21) International Application Number: PCT/GB93/01282

(22) International Filing Date: 16 June 1993 (16.06.93)

(30) Priority data:

9213376.8	24 June 1992 (24.06.92)	GB
9304057.4	1 March 1993 (01.03.93)	GB

(71) Applicant (*for all designated States except US*): SMITHKLINE BEECHAM PLC [GB/GB]; New Horizons Court, Brentford, Middlesex TW8 9EP (GB).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*) : SMITH, Richard, Anthony, Godwin [GB/GB]; DODD, Ian [GB/GB]; FREEMAN, Anne, Mary [GB/GB]; MOSSAKOWSKA, Danuta, Eva, Irena [GB/GB]; SmithKline Beecham Pharmaceuticals, Great Burgh, Yew Tree Bottom Road, Epsom, Surrey KT18 5XQ (GB).

(74) Agent: VALENTINE, Jill, Barbara; SmithKline Beecham, Corporate Patents, Great Burgh, Yew Tree Bottom Road, Epsom, Surrey KT18 5XQ (GB).

(81) Designated States: JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published

With international search report.

(54) Title: SOLUBLE CR1 DERIVATIVES

(57) Abstract

A soluble polypeptide comprising, in sequence, one to four short consensus repeats (SCR) selected from SCR 1,2,3 and 4 of long homologous repeat A(LHR-A) as the only structurally and functionally intact SCR domains of CR1 and including at least SCR3.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NE	Niger
BE	Belgium	GN	Guinea	NL	Netherlands
BF	Burkina Faso	GR	Greece	NO	Norway
BG	Bulgaria	HU	Hungary	NZ	New Zealand
BJ	Benin	IE	Ireland	PL	Poland
BR	Brazil	IT	Italy	PT	Portugal
BY	Belarus	JP	Japan	RO	Romania
CA	Canada	KP	Democratic People's Republic of Korea	RU	Russian Federation
CF	Central African Republic			SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovak Republic
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	UA	Ukraine
DE	Germany	MG	Madagascar	US	United States of America
DK	Denmark	ML	Mali	UZ	Uzbekistan
ES	Spain	MN	Mongolia	VN	Viet Nam
FI	Finland				

SOLUBLE CR1 DERIVATIVES.

The present invention relates to polypeptides and their use in the diagn sis and therapy of disorders involving complement activity and various inflammatory and immune disorders.

Constituting about 10% of the globulins in normal serum, the complement system is composed of many different proteins that are important in the immune system's response to foreign antigens. The complement system becomes activated when its primary components are cleaved and the products alone or with other proteins, activate additional complement proteins resulting in a proteolytic cascade. Activation of the complement system leads to a variety of responses including increased vascular permeability, chemotaxis of phagocytic cells, activation of inflammatory cells, opsonization of foreign particles, direct killing of cells and tissue damage. Activation of the complement system may be triggered by antigen-antibody complexes (the classical pathway) or, for example, by lipopolysaccharides present in cell walls of pathogenic bacteria (the alternative pathway).

Complement receptor type 1 (CR1) has been shown to be present on the membranes of erythrocytes, monocytes/macrophages, granulocytes, B cells, some T cells, splenic follicular dendritic cells, and glomerular podocytes. CR1 binds to the complement components C3b and C4b and has also been referred to as the C3b/C4b receptor. The str::ural organisation and primary sequence of one allotype of CR1 is known (Klickstein *et al.*, 1987, J. Exp. Med. 165:1095-1112, Klickstein *et al.*, 1988, J. Exp. Med. 168:1699-1717; Hourcade *et al.*,1988, J. Exp. Med. 168:1255-1270, WO 89/09220, WO 91/05047). It is composed of 30 short consensus repeats (SCRs) that each contain around 60-70 amino acids. In each SCR, around 29 of the average 65 amino acids are conserved. Each SCR has been proposed to form a three dimensional triple loop structure through disulphide linkages with the third and first and the fourth and second half-cystines in disulphide bonds. CR1 is further arranged as 4 long homologous repeats (LHRs) of 7 SCRs each. Following a leader sequence, the CR1 molecule consists of the N-terminal LHR-A, the next two repeats, LHR-B and LHR-C, and the most C-terminal LHR-D followed by 2 additional SCRs, a 25 residue putative transmembrane region and a 43 residue cytoplasmic tail.

Based on the mature CR1 molecule having a predicted N-terminal glutamine residue, hereinafter designated as residue 1, the first four SCR domains of LHR-A are defined herein as consisting f residues 2-58, 63-120, 125-191 and 197-252, respectively, of mature CR1.

Hourcade *et al.*,1988, J. Exp. Med. 168:1255-1270 observed an alternative polyadenylation site in the human CR1 transcriptional unit that was predicted to

produce a secreted form of CR1. The mRNA encoded by this truncated sequence comprises the first 8.5 SCRs of CR1, and encodes a protein of about 80 kDa which was proposed to include the C4b binding domain. When a cDNA corresponding to this truncated sequence was transfected into COS cells and expressed, it demonstrated the expected C4b binding activity but did not bind to C3b (Krych *et al.*,1989, FASEB J. 3:A368; Krych *et al.* Proc. Nat. Acad. Sci. 1991, 88, 4353-7). Krych *et al.*, also observed a mRNA similar to the predicted one in several human cell lines and postulated that such a truncated soluble form of CR1 with C4b binding activity may be synthesised in humans.

In addition, Makrides *et al.* (1992, J. Biol. Chem. 267 (34) 24754-61) have expressed SCR 1 + 2 and 1 + 2 + 3 + 4 of LHR-A as membrane-attached proteins in CHO cells.

Several soluble fragments of CR1 have also been generated via recombinant DNA procedures by eliminating the transmembrane region from the DNAs being expressed (WO 89/09220, WO 91/05047). The soluble CR1 fragments were functionally active, bound C3b and/or C4b and demonstrated Factor I cofactor activity depending upon the regions they contained. Such constructs inhibited *in vitro* complement-related functions such as neutrophil oxidative burst, complement mediated hemolysis, and C3a and C5a production. A particular soluble construct, sCR1/pBSCR1c, also demonstrated *in vivo* activity in a reversed passive Arthus reaction (WO 89/09220, WO 91/05047; Yeh *et al.*, 1991, J. Immunol. 146:250), suppressed post-ischemic myocardial inflammation and necrosis (WO 89/09220, WO 91/05047; Weisman *et al.*, Science, 1990, 249:146-1511; Dupe, R. *et al.* Thrombosis & Haemostasis (1991) 65(6) 695.) and extended survival rates following transplantation (Pruitt & Bollinger, 1991, J. Surg. Res 50:350; Pruitt *et al.*, 1991 Transplantation 52; 868). Furthermore, co-formulation of sCR1/pBSCR1c with p-anisoylated human plasminogen-streptokinase-activator complex (APSAC) resulted in similar anti-haemolytic activity as sCR1 alone, indicating that the combination of the complement inhibitor sCR1 with a thrombolytic agent was feasible (WO 91/05047).

Soluble polypeptides corresponding to part of CR1 have now been found to possess functional complement inhibitory, including anti-haemolytic, activity.

According to the present invention there is provided a soluble polypeptide comprising, in sequence, one to four short consensus repeats (SCR) selected from SCR 1, 2, 3 and 4 of l ng homologous repeat A (LHR-A) as the only structurally and functionally intact SCR domains of CR1 and including at least SCR3.

In preferred aspects, the polypeptide comprises, in sequence, SCR 1, 2, 3 and 4 of LHR-A or SCR 1, 2 and 3 of LHR-A as the only structurally and functionally

intact SCR domains of CR1.

It is to be understood that variations in the amino acid sequence of the polypeptide of the invention by way of addition, deletion or conservative substitution of residues, including allelic variations, in which the biological activity of the polypeptide is retained, are encompassed by the invention. Conservative substitution is understood to mean the retention of the charge and size characteristics of the amino acid side chain, for example arginine replaced by histidine.

In one aspect, the polypeptide of the invention may be represented symbolically as follows:

$$NH_2\text{-}V^1\text{-}SCR1\text{-}W^1\text{-}SCR2\text{-}X^1\text{-}SCR3\text{-}Y^1\text{-}OH \qquad (I)$$

in which SCR1 represents residues 2-58 of mature CR1, SCR2 represents residues 63-120 of mature CR1, SCR3 represents residues 125-191 of mature CR1, and V^1, W^1, X^1 and Y^1 represent bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

In a preferred embodiment of formula (I), W^1, X^1 and Y^1 represent residues 59-62, 121-124 and 192-196, respectively, of mature CR1 and V^1 represents residue 1 of mature CR1 optionally linked via its N-terminus to methionine.

In another aspect the polypeptide of the invention may be represented symbolically as follows:

$$NH_2\text{-}V^2\text{-}SCR1\text{-}W^2\text{-}SCR2\text{-}X^2\text{-}SCR3\text{-}Y^2\text{-}SCR4\text{-}Z^2OH \qquad (II)$$

in which SCR1, SCR2 and SCR3 are as hereinbefore defined, SCR4 represents residues 197-252 of mature CR1 and V^2, W^2, X^2, Y^2 and Z^2 represents bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

In preferred embodiments of formula (II), W^2, X^2, Y^2 and Z^2 represent residues 59-62, 121-124, 192-196, and residues 253 respectively, of mature CR1 and V^2 represents residue 1 of mature CR1 optionally linked via its N-terminus to methionine.

In one particular embodiment of formula (II) arginine 235 is replaced by histidine.

In the preferred embodiment of formula (II), residue 235 is arginine.

In one further aspect, the polypeptide of the invention may be represented symbolically as follows:

$$NH_2\text{-}X^3\text{-}SCR3\text{-}Y^3\text{-}OH \qquad (III)$$

in which SCR3 is as hereinbefore defined and X^3 and Y^3 represent bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

Figure 1

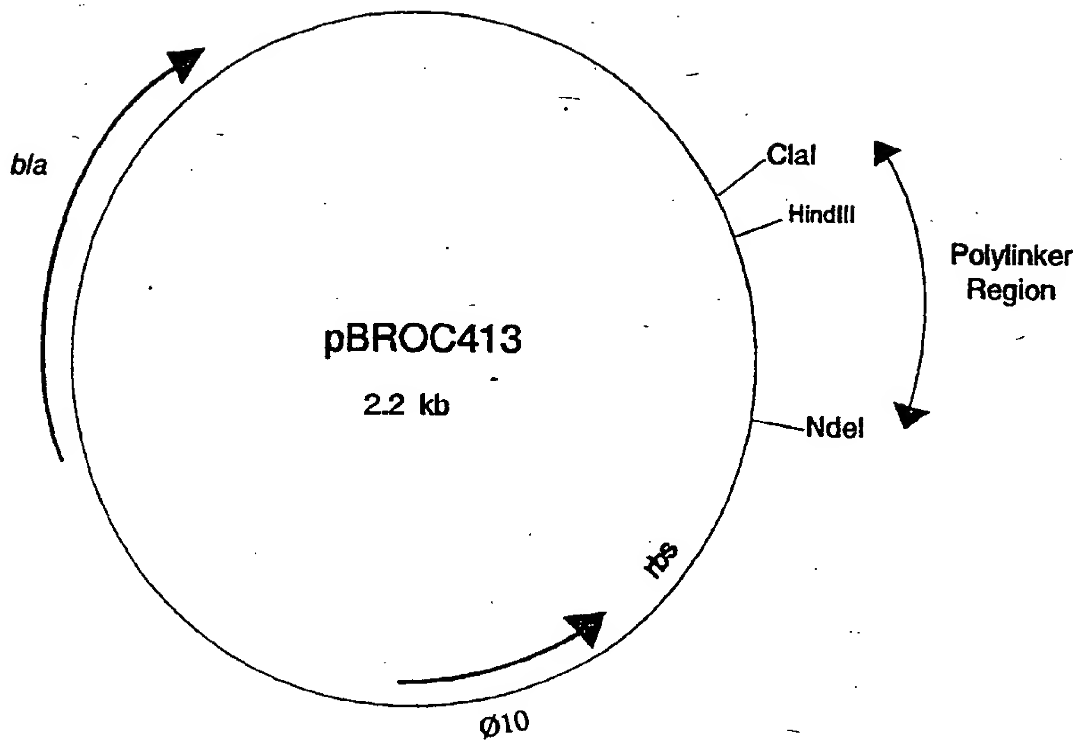

bla
pBROC413
2.2 kb
ClaI
HindIII
Polylinker
Region
NdeI
rbs
Ø10

Figure 2

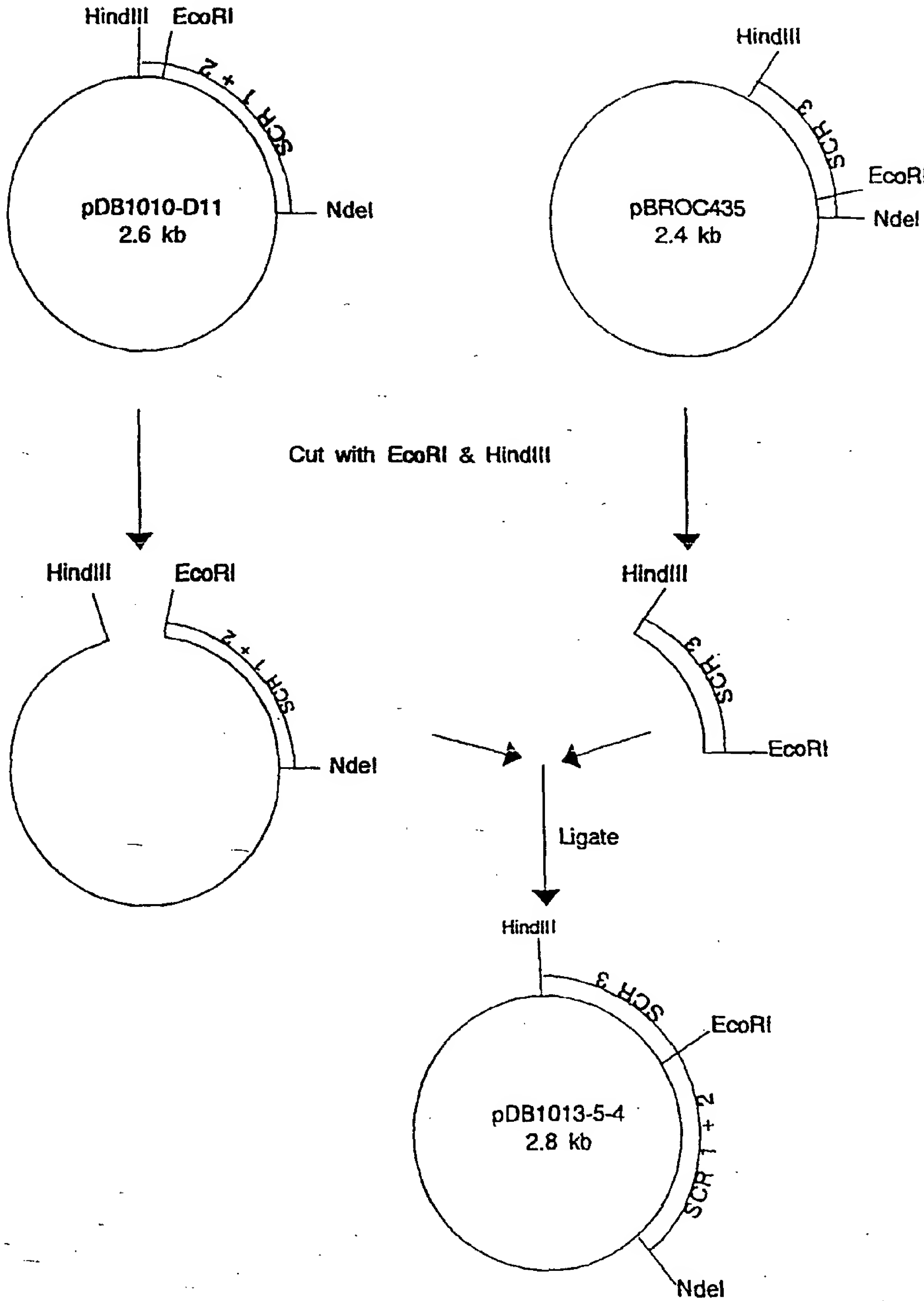
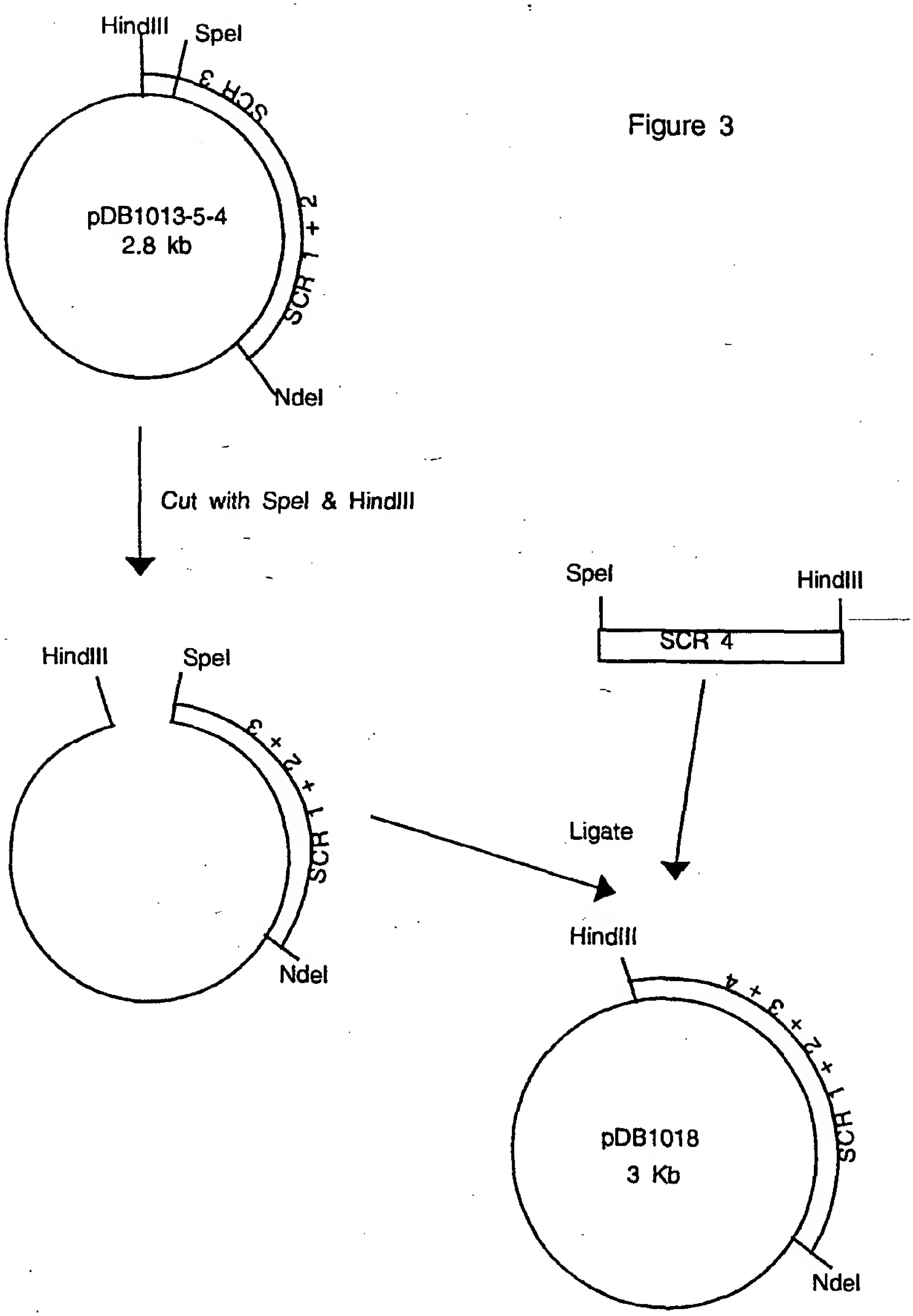

HindIII
EcoRI
SCR 1 + 2
NdeI
pDB1010-D11
2.6 kb
HindIII
SCR 3
EcoRI
NdeI
pBROC435
2.4 kb
Cut with EcoRI & HindIII
HindIII
EcoRI
SCR 1 + 2
NdeI
HindIII
SCR 3
EcoRI
Ligate
HindIII
SCR 3
EcoRI
SCR 1 + 2
NdeI
pDB1013-5-4
2.8 kb

Figure 3


HindIII
SpeI
SCR 3
SCR 1 + 2
pDB1013-5-4
2.8 kb
NdeI
Cut with SpeI & HindIII
HindIII
SpeI
SCR 1 + 2 + 3
NdeI
SpeI
HindIII
SCR 4
Ligate
HindIII
SCR 1 + 2 + 3 + 4
pDB1018
3 Kb
NdeI

INTERNATIONAL SEARCH REPORT

International Application No PCT/GB 93/01282

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[6]

According to International Patent Classification (IPC) or to both National Classification and IPC

Int.Cl. 5 C12N15/12; A61K37/02; C07K13/00; C12N1/21
//(C12N1/21,C12R1:19)

II. FIELDS SEARCHED

Minimum Documentation Searched[7]

Classification System	Classification Symbols
Int.Cl. 5	C07K ; C12N ; A61K

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched[8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT[9]

Category°	Citation of Document,[11] with indication, where appropriate, of the relevant passages[12]	Relevant to Claim No.[13]
A	WO,A,8 909 220 (THE JOHNS HOPKINS UNIVERSITY ET AL.) 5 October 1989 cited in the application see pages 32-38, chapters 5.5 and 5.6 see example 11.2.4.3 see figure 20 ---	1-21
A	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA vol. 88, no. 10, 15 May 1991, WASHINGTON US pages 4353 - 4357 KRYCH, M. ET AL. 'Sites within the complement C3b/C4b receptor important for the specificity of ligand binding' cited in the application see the whole document ---	1-21

-/--

° Special categories of cited documents :[10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

2

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
13 OCTOBER 1993	22. 10. 93
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	ANDRES S.M.

In a preferred embodiment of formula (III) X^3 represents amino acids 122-124 of mature CR1 optionally linked to methionine at its N-terminus and Y^4 represents amino acids 192-196 of mature CR1.

In another further aspect, the polypeptide of the invention may be represented symbolically as follows:

$$NH_2\text{-}X^4\text{-SCR3-}Y^4\text{-SCR4-}Z^4\text{-OH} \qquad \text{(IV)}$$

in which SCR3 and SCR4 are as hereinbefore defined and X^4, Y^4 and Z^4 represent bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

In a preferred embodiment of formula (IV) X^4 represents amino acids 122-124 of mature CR1 optionally linked to methionine at its N-terminus and Y^4 and Z^4 represent amino acids 192-196 and 253 respectively of mature CR1.

In a further aspect, the invention provides a process for preparing a CR1 polypeptide according to the invention which process comprises expressing DNA encoding said polypeptide in a recombinant host cell and recovering the product.

In particular, the process may comprise the steps of:

i) preparing a replicable expression vector capable, in a host cell, of expressing a DNA polymer comprising a nucleotide sequence that encodes said polypeptide;

ii) transforming a host cell with said vector;

iii) culturing said transformed host cell under conditions permitting expression of said DNA polymer to produce said polypeptide; and

iv) recovering said polypeptide.

The DNA polymer comprising a nucleotide sequence that encodes the polypeptide also forms part of the invention.

The process of the invention may be performed by conventional recombinant techniques such as described in Sambrook *et al.*, Molecular Cloning : A laboratory manual 2nd Edition. Cold Spring Harbor Laboratory Press (1989) and DNA Cloning vols I, II and III (D. M. Glover ed.. IRL Press Ltd).

The invention also provides a process for preparing the DNA polymer by the condensation of appropriate mono-, di- or oligomeric nucleotide units.

The preparation may be carried out chemically, enzymatically, or by a combination of the two methods, *in vitro* or *in vivo* as appropriate. Thus, the DNA polymer may be prepared by the enzymatic ligation of appropriate DNA fragments, by conventional methods such as those described by D. M. Roberts *et al.*, in Biochemistry 1985, 24, 5090-5098.

The DNA fragments may be obtained by digestion of DNA containing the required sequences of nucleotides with appropriate restriction enzymes, by chemical synthesis, by enzymatic polymerisation, or by a combination of these methods.

Digestion with restriction enzymes may be performed in an appropriate buffer at a temperature of 20°-70°C, generally in a volume of 50µl or less with 0.1-10µg DNA.

Enzymatic polymerisation of DNA may be carried out *in vitro* using a DNA polymerase such as DNA polymerase 1 (Klenow fragment) in an appropriate buffer containing the nucleoside triphosphates dATP, dCTP, dGTP and dTTP as required at a temperature of 10°-37°C, generally in a volume of 50µl or less.

Enzymatic ligation of DNA fragments may be carried out using a DNA ligase such as T4 DNA ligase in an appropriate buffer at a temperature of 4°C to 37°C, generally in a volume of 50µl or less.

The chemical synthesis of the DNA polymer or fragments may be carried out by conventional phosphotriester, phosphite or phosphoramidite chemistry, using solid phase techniques such as those described in 'Chemical and Enzymatic Synthesis of Gene Fragments - A Laboratory Manual' (ed. H.G. Gassen and A. Lang), Verlag Chemie, Weinheim (1982), or in other scientific publications, for example M.J.Gait, H.W.D. Matthes M. Singh, B.S. Sproat and R.C. Titmas, Nucleic Acids Research, 1982, 10, 6243; B.S. Sproat and W. Bannwarth, Tetrahedron Letters, 1983, 24, 5771; M.D. Matteucci and M.H. Caruthers, Tetrahedron Letters, 1980, 21, 719; M.D. Matteucci and M.H. Caruthers, Journal of the American Chemical Society, 1981, 103, 3185; S.P. Adams *et al.*, Journal of the American Chemical Society, 1983, 105, 661; N.D. Sinha, J. Biernat, J. McMannus and H. Koester, Nucleic Acids Research, 1984, 12, 4539; and H.W.D. Matthes *et al.*, EMBO Journal, 1984, 3, 801. Preferably an automated DNA synthesiser (for example, Applied Biosystems 381A Synthesiser) is employed.

The DNA polymer is preferably prepared by ligating two or more DNA molecules which together comprise a DNA sequence encoding the polypeptide.

The DNA molecules may be obtained by the digestion with suitable restriction enzymes of vectors carrying the required coding sequences.

The precise structure of the DNA molecules and the way in which they are obtained depends upon the structure of the desired product. The design of a suitable strategy for the construction of the DNA molecule coding for the polypeptide is a routine matter for the skilled worker in the art.

In particular, consideration may be given to the codon usage of the particular host cell. The codons may be optimised for high level expression in *E. coli* using the principles set out in Devereux *et al.*, (1984) Nucl. Acid Res., 12, 387.

The expression of the DNA polymer encoding the polypeptide in a recombinant host cell may be carried out by means of a replicable expression vector capable, in the host cell, of expressing the DNA polymer. The expression vector is

novel and also forms part of the invention.

The replicable expression vector may be prepared in accordance with the invention, by cleaving a vector compatible with the host cell to provide a linear DNA segment having an intact replicon, and combining said linear segment with one or more DNA molecules which, together with said linear segment, encode the polypeptide, under ligating conditions.

The ligation of the linear segment and more than one DNA molecule may be carried out simultaneously or sequentially as desired.

Thus, the DNA polymer may be preformed or formed during the construction of the vector, as desired. The choice of vector will be determined in part by the host cell, which may be prokaryotic, such as *E. coli*, or eukaryotic, such as mouse C127, mouse myeloma, chinese hamster ovary, fungi e.g. filamentous fungi or unicellular 'yeast' or an insect cell such as Drosophila. The host cell may also be in a transgenic animal. Suitable vectors include plasmids, bacteriophages, cosmids and recombinant viruses derived from, for example, baculoviruses or vaccinia.

The DNA polymer may be assembled into vectors designed for isolation of stable transformed mammalian cell lines expressing the fragment e.g. bovine papillomavirus vectors in mouse C127 cells, or amplified vectors in chinese hamster ovary cells (DNA Cloning Vol. II D.M. Glover ed. IRL Press 1985; Kaufman, R.J. *et al*.. Molecular and Cellular Biology 5, 1750-1759, 1985; Pavlakis G.N. and Hamer, D.H. Proceedings of the National Academy of Sciences (USA) 80, 397-401, 1983; Goeddel, D.V. *et al*.,European Patent Application No. 0093619, 1983).

The preparation of the replicable expression vector may be carried out conventionally with appropriate enzymes for restriction, polymerisation and ligation of the DNA, by procedures described in, for example, Sambrook *et al*.,cited above. Polymerisation and ligation may be performed as described above for the preparation of the DNA polymer. Digestion with restriction enzymes may be performed in an appropriate buffer at a temperature of 20°-70°C, generally in a volume of 50μl or less with 0.1-10μg DNA.

The recombinant host cell is prepared, in accordance with the invention, by transforming a host cell with a replicable expression vector of the invention under transforming conditions. Suitable transforming conditions are conventional and are described in, for example, Sambrook *et al*., cited above, or "DNA Cloning" Vol. II, D.M. Glover ed., IRL Press Ltd, 1985.

The choice of transforming conditions is determined by the host cell. Thus, a bacterial host such as *E.coli*, may be treated with a solution of $CaCl_2$ (Cohen *et al*.,Proc. Nat. Acad. Sci., 1973, <u>69</u>, 2110) or with a solution comprising a mixture of RbCl, $MnCl_2$, potassium acetate and glycerol, and then with 3-[N-morpholino]-

propane-sulphonic acid, RbCl and glycerol or by electroporation as for example described by Bio-Rad Laboratories, Richmond, California, USA, manufacturers of an electroporator. Mammalian cells in culture may be transformed by calcium co-precipitation of the vector DNA onto the cells or by using cationic liposomes.

The invention also extends to a host cell transformed with a replicable expression vector of the invention.

Culturing the transformed host cell under conditions permitting expression of the DNA polymer is carried out conventionally, as described in, for example, Sambrook *et al.*, and "DNA Cloning" cited above. Thus, preferably the cell is supplied with nutrient and cultured at a temperature below 45°C.

The protein product is recovered by conventional methods according to the host cell. Thus, where the host cell is bacterial such as *E. coli* and the protein is expressed intracellularly, it may be lysed physically, chemically or enzymatically and the protein product isolated from the resulting lysate. Where the host cell is mammalian, the product is usually isolated from the nutrient medium.

Where the host cell is bacterial, such as *E. coli*, the product obtained from the culture may require folding for optimum functional activity. This is most likely if the protein is expressed as inclusion bodies. There are a number of aspects of the isolation and folding process that are regarded as important. In particular, the polypeptide is preferably partially purified before folding, in order to minimise formation of aggregates with contaminating proteins and minimise misfolding of the polypeptide. Thus, the removal of contaminating *E. coli* proteins by specifically isolating the inclusion bodies and the subsequent additional purification prior to folding are important aspects of the procedure.

The folding process is carried out in such a way as to minimise [illegible]ion of intermediate-folded states of the polypeptide. Thus, careful consideration needs to be given to, among others, the salt type and concentration, temperature, protein concentration, redox buffer concentrations and duration of folding. The exact condition for any given polypeptide generally cannot be predicted and must be determined by experiment.

There are numerous methods available for the folding of proteins from inclusion bodies and these are known to the skilled worker in this field. The methods generally involve breaking all the disulphide bonds in the inclusion body, for example with 50mM 2-mercaptoethanol, in the presence of a high concentration of denaturant such as 8M urea or 6M guanidine hydrochloride. The next step is to remove these agents to allow folding of the proteins to occur. Formation of the disulphide bridges requires an oxidising environment and this may be provided in a number of ways, for example by air, or by incorporating a suitable redox system, for example a mixture of

reduced and oxidised glutathione.

Preferably, the inclusion body is solubilised using 8M urea, in the presence of mercaptoethanol, and protein is folded, after initial removal of contaminating proteins, by addition of cold buffer. A preferred buffer is 20mM ethanolamine containing 1mM reduced glutathione and 0.5mM oxidised glutathione. The folding is preferably carried out at a temperature in the range 1 to $5^{0}C$ over a period of 1 to 4 days.

If any precipitation or aggregation is observed, the aggregated protein can be removed in a number of ways, for example by centrifugation or by treatment with precipitants such as ammonium sulphate. Where either of these procedures are adopted, monomeric polypeptide is the major soluble product.

If the bacterial cell secretes the protein, folding is not usually necessary.

The polypeptide of this invention is useful in the treatment or diagnosis of many complement-mediated or complement-related diseases and disorders including, but not limited to, those listed below.

Disease and Disorders Involving Complement

Neurological Disorders

multiple sclerosis
stroke
Guillain Barré Syndrome
traumatic brain injury
Parkinson's disease
allergic encephalitis
Alzheimer's disease

Disorders of Inappropriate or Undesirable Complement Activation

haemodialysis complications
hyperacute allograft rejection
xenograft rejection
corneal graft rejection
interleukin-2 induced toxicity during IL-2 therapy
paroxysmal nocturnal haemoglobinuria

Inflammatory Dis rders

inflammation of autoimmune diseases
Crohn's Disease
adult respiratory distress syndrome

thermal injury including burns or frostbite
uveitis
psoriasis
asthma
acute pancreatitis

Post-Ischemic Reperfusion Conditions
myocardial infarction
balloon angioplasty
atherosclerosis (cholesterol-induced) & restenosis
hypertension
post-pump syndrome in cardiopulmonary bypass or renal haemodialysis
renal ischemia
intestinal ischaemia

Infectious Diseases or Sepsis
multiple organ failure
septic shock

Immune Complex Disorders and Autoimmune Diseases
rheumatoid arthritis
systemic lupus erythematosus (SLE)
SLE nephritis
proliferative nephritis
glomerulonephritis
haemolytic anemia
myasthenia gravis

Reproductive Disorders
antibody- or complement-mediated infertility

Wound Healing

The present invention is also directed to a pharmaceutical composition comprising a therapeutically effective amount of a polypeptide, as above, and a pharmaceutically acceptable carrier or excipient.

The present invention also provides a method of treating a disease or disorder associated with inflammation or inappropriate complement activation comprising

administering to a subject in need of such treatment a therapeutically effective amount of a polypeptide of this invention.

In the above methods, the subject is preferably a human.

An effective amount of the polypeptide for the treatment of a disease or disorder is in the dose range of 0.01-100mg/kg; preferably 0.1mg-10mg/kg.

For administration, the polypeptide should be formulated into an appropriate pharmaceutical or therapeutic composition. Such a composition typically contains a therapeutically active amount of the polypeptide and a pharmaceutically acceptable excipient or carrier such as saline, buffered saline, dextrose, or water. Compositions may also comprise specific stabilising agents such as sugars, including mannose and mannitol, and local anaesthetics for injectable compositions, including, for example, lidocaine.

Further provided is the use of a polypeptide of this invention in the manufacture of a medicament for the treatment of a disease or disorder associated with inflammation or inappropriate complement activation.

In order to inhibit complement activation and, at the same time, provide thrombolytic therapy, the present invention provides compositions which further comprise a therapeutically active amount of a thrombolytic agent. An effective amount of a thrombolytic agent is in the dose range of 0.01-10mg/kg; preferably 0.1-5mg/kg. Preferred thrombolytic agents include, but are not limited to, streptokinase, human tissue type plasminogen activator and urokinase molecules and derivatives, fragments or conjugates thereof. The thrombolytic agents may comprise one or more chains that may be fused or reversibly linked to other agents to form hybrid molecules (EP-A-0297882 and EP 155387), such as, for example, urokinase linked to plasmin (EP-A-0152736), a fibrinolytic enzyme linked to a water-soluble polymer (EP-A-0183503). The thrombolytic agents may also comprise muteins of plasminogen activators (EP-A-0207589). In a preferred embodiment, the thrombolytic agent may comprise a reversibly blocked *in vitro* fibrinolytic enzyme as described in U.S. Patent No. 4,285,932. A most preferred enzyme is a p-anisoyl plasminogen-streptokinase activator complex as described in U.S. Patent No. 4,808,405, and marketed by SmithKline Beecham Pharmaceuticals under the Trademark EMINASE (generic name anistreplase, also referred to as APSAC; Monk *et al.*,1987, Drugs 34:25-49).

Routes of administration for the individual or combined therapeutic compositions of the present invention include standard routes, such as, for example, intravenous infusion or bolus injecti n. Active complement inhibitors and thrombolytic agents may be administered together or sequentially, in any order.

The present invention also provides a method for treating a thrombotic condition, in particular acute myocardial infarction, in a human or non-human animal.

This method comprises administering to a human or animal in need of this treatment an effective amount of a polypeptide according to this invention and an effective amount of a thrombolytic agent.

Also provided is the use of a polypeptide of this invention and a thrombolytic agent in the manufacture of a medicament for the treatment of a thrombotic condition in a human or animal. Such methods and uses may be carried out as described in WO 91/05047.

This invention further provides a method for treating adult respiratory distress syndrome (ARDS) in a human or non-human animal. This method comprises administering to the patient an effective amount of a polypeptide according to this invention.

The invention also provides a method of delaying hyperacute allograft or hyperacute xenograft rejection in a human or non-human animal which receives a transplant by administering an effective amount of a polypeptide according to this invention. Such administration may be to the patient or by application to the transplant prior to implantation.

The invention yet further provides a method of treating wounds in a human or non-human animal by administering by either topical or parenteral e.g. intravenous routes, an effective amount of a polypeptide according to this invention.

GENERAL METHODS USED IN EXAMPLES

(i) **DNA cleavage**

Cleavage of DNA by restriction endonucleases was carried out according to the manufacturer's instructions using supplied buffers. Double digests were carried out simultaneously if the buffer conditions were suitable for both enzymes. Otherwise double digests were carried out sequentially where enzyme requiring the lowest salt concentration was added first to the digest. Once that digest was complete the salt concentration was altered and the second enzyme added.

(ii) **Production of blunt ended DNA fragments**

The recessed 3' termini of DNA fragments were filled in using the Klenow fragment of DNA polymerase I as described in Sambrook *et al* (1989).

(iii) **DNA purification/ concentration and analysis**

Removal of protein contaminants, nucleosides was with phenol/$CHCl_3$ followed by precipitation with ethanol. DNA was analysed on horizontal agarose gel electrophoresis; both methods are described in Sambrook *et al* (1989).

(iv) **DNA fragment isolation**

1. **DNA purification n DEAE NA45 membranes**

DNA fragments were purified from agarose gels by making an incision in the

agarose above and just below the required DNA fragment. NA45 membranes from Schleicher & Schuell (Anderman, Great Britain) that had been soaked in TE (10 mM Tris pH 8.0, 1 mM EDTA) were inserted into the incisions and current reapplied to the gel until the DNA fragment was trapped on the lower membrane; higher molecular weight DNA was trapped on the upper membrane. The lower membrane was removed from the gel and the DNA eluted into 0.05 M arginine/1 M NaCl at 70°C for 2 hours. The DNA was then concentrated by ethanol precipitation as described in Sambrook *et al* (1989).

2. **Electroelution**

DNA fragments were excised from agarose gels and DNA extracted by electroelution using the Unidirectional Electroeluter (IBI Ltd.,Cambridge, England) according to the manufacturer's instructions.

3. **Gel purification**

DNA fragments were excised from agarose gels and DNA extracted using the QIAEX gel extraction kit according to the manufacturers instructions (QIAGEN Inc., USA).

(v) **Plasmid preparation**

Large scale plasmid preparation of plasmid DNA was carried out using CsCl as described in Sambrook *et al* (1989) or using Magic Maxipreps (Promega Corporation, Madison, USA) according to the manufacturers instructions. Mini-plasmid preparations were carried out using either the alkaline lysis method described in Sambrook *et al* (1989) or Magic Minipreps (Promega Corporation, Madison, USA) according to the manufacturer's instructions.

(vi) **Introduction of plasmid DNA into *E. coli***

1. Plasmids were transformed into *E. coli* HB101 or *E. coli* BL21 (DE3) (Studier and Moffat, 1986) that had been made competent using calcium chloride as described in Sambrook *et al* (1989).

2. Alternatively plasmids were introduced into *E.coli* DH1 (Low,1968) or *E.coli* BL21 (DE3) by electroporation using the Gene Pulsar and Pulse Controller of Bio-Rad (Bio-Rad Laboratories, Richmond, California,USA) according to the manufacturer's instructions.

(vii) **Kinasing of oligonucleotides**

Oligonucleotides or annealed oligonucleotides possessing 5' overhangs were kinased using T_4 polynucleotide kinase as described in Sambrook *et al* (1989).

(viii) **Annealing and ligation of oligonucleotides**

Oligonucleotides were annealed together by mixing generally equimolar concentrations of the complementary oligonucleotides in 10 mM Tris pH 8.5, 5 mM $MgCl_2$ and placing at 100°C for 5 minutes and then cooling very slowly to room

temperature. Annealed oligonucleotides with sticky ends were ligated to vector or other oligonucleotides containing complementary sticky ends using T_4 DNA ligase as described in Sambrook *et al* (1989).

(ix) **PCR (Polymerase Chain Reaction) amplification of DNA**

DNA fragments from ligation reactions or DNA fragments excised and purified from agarose gels were amplified by PCR from two primers complementary to the 5' ends of the DNA fragment. Approximately 0.1 - 1μg of ligation reaction or the purified DNA from the agarose gel was mixed in 10 mM Tris pH 8.3 (at 25°C), 50 mM KCl, 0.1% gelatin; $MgCl_2$ concentrations were varied from 1.5 mM to 6 mM to find a suitable concentration for each reaction. Both primers were added to a final concentration of 2 μM; each dNTP was added to a final concentration of 0.2 mM. The final reaction volume was either 75 μl or 100 μl, which was overlayed with mineral oil to prevent evaporation. Thermal cycling was then started on a thermal cycler eg. Hybaid Thermal reactor, and a typical example of conditions used was 94°C 7 mins, 45°C 2 mins, hold at 45°C for less than 5 min., and then add 5 units of Taq DNA polymerase (purchased from a commercial source, e.g. Gibco). The DNA fragment was amplified by cycling the temperature at 72°C 2 mins, 94°C 1 min and 45°C 2 min a total of 35 times.

(x) **DNA sequencing using the double stranded method**

Sequencing was carried out using "Sequenase™" (United States Biochemical Corporation) essentially as described in the manufacturer's instructions.

(xi) **DNA sequence analysis and manipulation**

Analysis of sequences were carried out on a digital VAX computer using the GCG package of programmes as described in Devereux *et al* (1984).

(xii) **Production of oligonucleotides**

1. Oligonucleotides were synthesised using a Gene Assembler Plus (Pharmacia LKB Biotechnology, Milton Keynes, England) or a 381A Synthesiser (Applied BioSystems) according to the manufacturer's instructions.

2. Oligonucleotide purification was carried out either using MonoQ as recommended by Pharmacia or by UV shadowing where recovery of synthetic oligonucleotides was by electrophoresis through a denaturing polyacrylamide gel. The oligonucleotides were loaded onto a 12% acrylamide/7M urea gel and run at 1500V until the oligonucleotide had migrated approximately two thirds of the length of the gel. The DNA was visualised using a hand-held, long-wavelength ultraviolet lamp; and the DNA bands excised. The oligonucleotide was recovered using Sep-Pak C18 reverse phase columns (Waters) as described in Sambrook et al (1989).

(xiii) **Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS PAGE)**

SDS PAGE was carried out generally using the Novex system (British Biotechnology) according to the manufacturer's instructions. Prepacked gels of acrylamide concentrations 14%, 16%, 4 - 20% or 10 - 27% were the ones most frequently used. Samples for electrophoresis, including protein molecular weight standards (LMW Kit, Pharmacia) were usually diluted in 1%(w/v)SDS - containing buffer (with or without 5%(v/v) 2-mercaptoethanol), and left at room temperature for about 0.5 to 1 h before application to the gel.

(xiv) **Alteration of codon usage**

The non random use of synonymous codons has been demonstrated in *E. coli* and there is some evidence to support the belief that protein production from genes containing non-optimal or minor codons (particularly at the 5' end of the gene) is less efficient than that from genes with no such codons (e.g. Chen an Inouye, 1990). A codon usage table compiled from genes highly expressed in *E. coli* (supplied as part of the GCG sequence analysis software package, Devereux *et al*, (1984)) was used to determine the optimal codons for expression in *E. coli*. All of the first 30 codons of all constructs (where compatible with restriction enzyme sites) were optimised for high level expression. The codons for the seven amino acids: arg, gly, ile, leu, pro, ser, ala were optimised (where compatible with restriction enzyme sites) throughout the coding sequence.

(xv) **Construction of vector pBROC413**

The plasmid pT7-7 (Tabor, 1990) contains DNA corresponding to nucleotides 2065-4362 of pBR322 and like pBR322 can be mobilized by a conjugative plasmid in the presence of a third plasmid ColK. A mobility protein encoded by ColK acts on the *nic* site at nucleotide 2254 of pBR322 initiating mobilization from this point. pT7-7 was digested with *Lsp*I and *Bgl*II and the protruding 5' ends filled in with the Klenow fragment of DNA Polymerase I. The plasmid DNA fragment was purified by agarose gel electrophoresis, the blunt ends ligated together and transformed into *E. coli* DH1 by electroporation. The resultant plasmid pBROC413 (Fig.1) was identified by restriction enzyme analysis of plasmid DNA.

The deletion in pBROC413 from the *Lsp*I site immediately upstream of the ø10 promoter to the *Bgl*II site at nucleotide 434 of pT7-7 deletes the DNA corresponding to nucleotides 2065-2297 of pBR322. The *nic* site and adjacent sequences are therefore deleted making pBROC413 non mobilizable.

(xvi) **Haemolytic assay**

The anti-haemolytic activity of polypeptides was assessed by measuring the inhibition of complement mediated lysis of sheep erythrocytes sensitised with rabbit antibodies (obtained from Diamedix Corporation, Miami, USA). Human serum diluted 1:125 in 0.1 M Hepes/ 0.15 M NaCl pH 7.4 buffer was the source of

complement and was prepared from a pool of volunteers essentially as described in (Dacie & Lewis, 1975). Briefly, blood was warmed to 37°C for 5 minutes, the clot removed and the remaining serum clarified by centrifugation. The serum fraction was split into small aliquots and stored at -196°C. Aliquots were thawed as required and diluted in the Hepes buffer immediately before use.

Inhibition of complement-mediated lysis of sensitised sheep erythrocytes was measured using a standard haemolytic assay using a v-bottom microtitre plate format as follows.

50 μl of a range of concentrations (0.01-100μg/ml but typically 0.05 - 25 μg/ml) of test protein diluted in Hepes buffer were incubated with 50 μl of the diluted serum for 15 minutes at 37°C. 100 μl of prewarmed sensitised sheep erythrocytes were added for 1 hour at 37°C in a final reaction volume of 200 μl. Samples were spun at 300g at 4°C for 15 minutes before transferring 150 μl of supernatant to flat bottomed microtitre plates and determining the absorption at 410 nm, which reflects the amount of lysis in each test solution. Maximum lysis was determined by incubating serum with erythrocytes in the absence of any inhibitor from which the proportion of background lysis had been subtracted (determined by incubating erythrocytes with buffer). The background lysis by inhibitor was assessed by incubating inhibitor with erythrocytes and then subtracting that from test samples. Inhibition was expressed as a fraction of the total cell lysis such that IH50 represents the concentration of inhibitor required to give 50% inhibition of lysis.

(xvii) **C3a RIA Assay**

Activation of complement pathways can be followed by measuring the release of the anaphylatoxin, C3a and its breakdown product C3a des Arg. Both products can be measured using a competitive radio-immuno assay purchased from Amersham International plc, U. K., (human complement C3a des Arg [^{125}I]assay, code RPA 518).

(a) **Alternative pathway activation by Zymosan A**

The alternative pathway of complement was activated with zymosan A, a complex carbohydrate from yeast (Sigma, catalogue number Z-4250). Zymosan A was made 50 mg/ml in Hepes buffer (0.1M Hepes/0.15M NaCl pH 7.4) or in PBS (50 mM sodium phosphate/0.1 M NaCl pH 7.4) and vortexed until a fine suspension had formed. Serum (prepared as described for the haemolytic assay; Method xvi) was preincubated with different concentrations of complement inhibitor diluted in Hepes buffer for 15 mins at 37°C using the volumes given below. Zymosan A was then vortexed for a few seconds each time before addition to the samples after which samples were incubated for a further 30 mins at 37°C. The zymosan A was then spun down at approximately 11,000g for 30 seconds at ambient temperature.

Typically 100 μl of supernatant were added to an equal volume of precipitating solution provided in the kit and the subsequent supernatant assayed as described in the technical bulletin supplied by Amersham with the C3a des Arg assay RIA kit. Each sample was assayed in duplicate and useful dilutions of the supernatant, to ensure that sample counts were on the standard curve, were found to be 1/50 - 1/100. EDTA or Futhan were not used in any solutions or tubes as suggested in the technical bulletin.

Each sample was counted for 1 minute on an LKB-Wallac 1272 Clinigamma. Data was processed using the RiaCalc program for RIA assays as supplied with the Clinigamma. The data was computed essentially as described in the Amersham technical bulletin with the standrd curve constructed by non-linear regression fit to the data.

The miniaturised assay was carried out essentially as described above but using smaller total volumes for the activation of serum.

Volumes of samples added

	serum	inhibitor	Zymosan A
Normal Assay	79 μl	20 μl	21 μl
Miniaturised Assay	26.3 μl	6.7 μl	7 μl

In the minaturised assay, after activation, typically 25 μl of the sample were precipitated. The assay kit reagent additions were reduced from 50 μl to 10 μl which enabled the assay to be carried out in a U-bottom microtitre plate containing separate detachable wells. The assay was then carried out as described in the technical bulletin using the adjusted volumes until the last dilution in isotonic saline. In this instance 200 μl of saline were added and the plate spun at approximately 2500g for 12 mins at 4°C. The supernatants from each well were carefully removed by aspiration and the precipitate was washed with a further 300 μl of isotonic saline. The plate was then spun again at about 2500g for 5 mins, 4°C and the supernatant was discarded. Wells were then counted for 10 mins each on the Clinigamma. The data was processed as above.

To determine the % inhibition of maximum activation at each inhibitor concentration, a number of controls were carried out with each experiment. These included maximum activation (A) *i.e.* serum + zymosan A only, background activation (B) *i.e.* serum + buffer only, and background activation in the presence of inhibitor (C) *i.e.* serum + inhibitor only. The background activation was generally subtracted from the maximum activation. Similarly the background activation in the presence of inhibitor was subtracted from the value of activated serum in the presence of inhibitor. These values could then be used to determine the % inhibition at each

inhibitor concentration. using the following formula:

$$1 - \frac{(D - C)}{(A - B)} \times 100$$

where D is the value of activation of serum in the presence of inhibitor and zymosan A. The IC50 is defined as the concentration of inhibitor required to reduce maximum activation by 50%.

(b) **Classical pathway activation by heat aggregated IgG**

Activation of the classical pathway by IgG was performed as follows. Human γ-globulin (Sigma, catalogue number G-4386) was made 14 mg/ml in 0.1 M Hepes/0.15 M NaCl pH 7.4 and heated at 60°C for 1 hour. Samples of heat aggregated IgG were then stored as small aliquots at -80°C until required. Serum was activated using heat aggregated IgG using the same volumes as described for the zymosan A normal or miniaturised assay. Preincubation of inhibitor with serum was for 15 mins at 37°C followed by addition of the heat aggregated IgG. Incubation was continued for a further 45 mins at 37°C. The samples were then assayed directly for C3a levels using either the normal or miniaturised assay.

(xviii) **C5a RIA Assay**

Activation of complement pathways can be followed by measuring the release of the anaphylatoxin C5a and its breakdown product C5a des Arg. Both products can be measured using a competitive radio-immuno assay purchased from Amersham International plc, U. K., (human complement C5a des Arg [^{125}I]assay, code RPA 520).

The alternative pathway of complement was activated with zymosan A, as described for the C3a RIA assay (Method (xvii)). The assay was carried out in the miniaturised form as described for the C3a assay using the reagents provided in the C5a des-Arg RIA kit.

References in the Examples to amino acid numbering relate to the corresponding residues of mature CR1 protein.

Example 1 Construction of plasmid pDB1010-D11 encoding SCR 1 + 2

General points

A DNA sequence for SCR 1 + 2 corresponding to amino acid 1 and ending at amino acid 124 of mature human complement receptor 1 was designed such that the 5' end of the gene contained an *Nde*I site. This site comprises an ATG codon to give the initiating methionine required for the start of mRNA translation and places the gene an optimum distance from the Shine-Dalgarno ribosome binding sequence of pBROC413. The 3' end of the gene finished on two stop codons followed by a

*Hin*dIII site.

Restriction endonucleases that do not cut pBROC413 and that were commercially available were identified. The sequences of the restriction sites recognised by the endonucleases were translated into all three reading frames. The sites that contained rarely used codons for *E. coli* expression were discarded. The remaining sites were matched with the DNA coding for SCR 1 + 2. If the restriction site could be fitted into the DNA sequence so as to preserve the coding sequence and not add a rarely used codon, the DNA sequence was altered to include this restriction site. 10 unique restriction sites were so identified and incorporated. To enable intracellular expression of protein in *E. coli*, an ATG codon was added to the 5' end of the gene immediately preceding the codon for the first amino acid of mature CR-1. The codon ATG is part of the *Nde*I restriction site which can be used for cloning into vectors such as pBROC413. The codon corresponding to proline 124 of mature CR-1 has been changed to one encoding glutamine, which also encompasses an *Eco*RI site.

(a) **Construction of plasmid**

Oligonucleotides coding for SCR 1 + 2 (Table 1; 1 - 8) were synthesised as 4 complementary pairs of 87 - 101 mers that could be ligated in a unique fashion via complementary 8 bp overhangs between the pairs of oligonucleotides. The four complementary pairs of oligonucleotides were designated Pair A (oligos 1+2), Pair B (oligos 3+4), Pair C (oligos 5+6) and Pair D (oligos 7+8). Pair A which corresponded to the 5' end of the gene contained an *Nde*I restriction site overhang and Pair D contained a *Hin*dIII restriction site overhang at the 3' end. All oligonucleotides apart from 1 and 2 were purified on Pharmacia Mono Q columns prior to use. Oligonucleotide 2 of pair A and oligonucleotide 7 of pair D were kinased before annealing with their unkinased complementary oligonucleotides 1 and 8 respectively. Oligonucleotides pairs B and C were annealed first and then kinased. The kinased oligonuleotide pairs were ligated Pair A (approx. 0.1μg) to Pair B (approx. 0.2μg) and Pair C (approx. 2μg) to Pair D (approx. 4μg). The ligated oligonucleotides (A+B) were in turn ligated to (C+D) to form the gene coding for SCR 1 + 2.

The DNA coding for SCR 1 + 2 was amplified by PCR using two oligonucleotides (Table 1; 15 and 16) complementary to the two strands of DNA. Both oligonucleotides contained 5' unmatched ends that contained 6 bp of random sequence followed by the sequence of either *Nde*I or *Hin*dIII restriction sites followed by 18 bp complementary to the gene. Following PCR, a band of approximately 400 bp was visualised on horizontal agarose gel electrophoresis, which was excised and purified on DEAE NA45 membranes. The DNA was then cut with *Nde*I and *Hin*dIII

before ligating into pBROC413 that had been cut with the same enzymes. The vector was transformed into *E. coli* HB101 made competent with calcium chloride. Mini-plasmid preparations were made and the plasmid DNA was analysed by digestion with *Nde*I and *Hin*dIII. Plasmids containing the correct sized insert, were further subjected to restriction mapping with *Eco*RI, *Hpa*I, *Kpn*I and *Sma*I. The plasmids that displayed the correct restriction maps were analysed by DNA sequencing of both strands across the gene coding for SCR 1 + 2. Plasmid pDB1010-D11 was identified as having the correct sequence across the gene coding for SCR 1 + 2.

Example 2 Construction, expression, purification, folding and formulation of MQ1 -> K196 of CR-1 (SCR 1+2+3)

General Points

The DNA coding for SCR 1+2+3 was constructed by ligating DNA coding for SCR 1+2 (Example 1a) to DNA encoding SCR 3.

General points relating to SCR 3 are presented in Example 9.

The SCR 3 coding unit corresponding to amino acid 122 and ending at amino acid 196 of mature CR1, was designed such that 5' end of the unit contained the *Eco*RI site at the junction of SCR's 2 & 3 as well as an *Nde*I site 5' to the *Eco*RI site. The 3' end of the unit finished on two stop codons followed by a *Hin*dIII site. The plasmids containing the SCR 3 coding unit and the SCR 1+2 coding unit were digested with *Eco*RI and *Hin*dIII. The SCR 3 coding unit was isolated and inserted downstream of the SCR 1+2 coding unit in the *Eco*RI/*Hin*dIII-cut SCR 1+2 - containing plasmid, to give a plasmid containing the SCR 1+2+3 coding unit, which corresponds to amino acids 1 to 196 of mature CR1. The addition of the SCR 3 coding unit through the *Eco*RI site, converts the codon corresponding to a glutamine at position 124 back to the authentic amino acid (proline) that is found in CR1.

(a) **Construction of plasmid pDB1013-5-4 encoding SCR 1 + 2 + 3**

Three pairs of oligonucleotides (Table 1; 9 - 14) encompassing the SCR 3 coding sequence were synthesised. The oligonucleotides were first annealed as pairs (9, 10; 11,12; 13,14) and the middle pair kinased thus allowing the three pairs to be ligated together via 8 base pair overlapping sequences. The 5' end of this molecule was designed to be complementary to *Nde*I digested DNA and the 3' end to *Hin*dIII digested DNA. This enabled the trimer to be cloned into *Nde*I/*Hin*dIII digested pBROC413 generating pBROC435 (Fig.2). The identity of pBROC435 was checked by restriction enzyme analysis and confirmed by DNA sequencing.

Plasmid DNA from pBROC435 and pDB1010-D11 (Example 1) were both cut with *Eco*RI and *Hin*dIII; the *Eco*RI/*Hin*dIII band of pBROC435 coding for SCR 3

was purified on an DEAE NA45 membrane as was the cut vector pDB1010-D11. The SCR 3 coding unit was then ligated into pDB1010-D11 to generate pDB1013-5 which was then transformed into calcium chloride competent *E. coli* HB101. The resulting colonies were analysed by mini-plasmid preparation of DNA followed by restriction mapping. One of the colonies, termed pDB1013-5-4 (Fig.2), contained the SCR 1+2+3 coding unit. This plasmid was then analysed for expression of the gene product.

(b) **Expression of SCR 1 + 2 + 3**

pDB1013-5-4 was transformed into calcium chloride competent *E.coli* BL21(DE3) and resulting colonies analysed by restriction digestion of mini-plasmid DNA preparations. Single colonies were inoculated into universals containing 10 ml of L broth or NCYZM medium and 50 µg/ml ampicillin and allowed to grow overnight at 37°C, 220 r.p.m. The overnight cultures (typically 5ml) were used to inoculate each of 2 L conical flasks containing 500 mls of NCYZM medium, 150 µg/ml ampicillin; cultures were grown at 37°C, 220 r.p.m. until A_{600} was 0.5 absorbance units. Cultures were induced with 1 mM isopropylthioβ-D-galactoside (IPTG) and allowed to grow a further 3 hours under the same conditions. The cultures were centrifuged (approx. 8000g/10 min) and the supernatants discarded. The cell pellets were stored at -40°C. L broth was 1% (w/v) Bactotryptone, 0.5% (w/v) Bactoyeastextract, 0.5% (w/v) NaCl. NCYZM media was L-broth containing 0.1% (w/v) casamino acids and 0.2%(w/v) $MgSO_4.7H_2O$, pH 7.0.

(c) **Isolation of solubilised inclusion bodies**

Frozen cell pellet of *E.coli* BL21 DE3 (pDB1013-5-4) (1 litre culture) prepared in a similar way to that described in Example 2b. was allowed to thaw at 4°C for 2 h and was then resuspended in 50 mM Tris/50 mM NaCl/1 mM EDTA/0.1 mM PMSF pH 8.0 (33 ml). The suspension was transferred to a 100 ml glass beaker and sonicated (Heat Systems - Ultrasonics W380; 70 Watts, 50 x 50% pulse, pulse time = 5 sec.). The sonicate was immediately centrifuged (6000g/4°C/10 min) and the supernatant was discarded. The pellet, containing the inclusion bodies, was resuspended in 20 mM Tris/8 M urea/50 mM 2-mercaptoethanol/1 mM EDTA/0.1 mM PMSF pH 8.5 (100 ml) and left static at room temperature (approx. 23°C) for 1 h. The resulting solution was centrifuged (approx. 2000 g at 4°C for 10 min) to remove any material that had failed to solubilise. The supernatant of this spin was retained at -40°C as the solubilised inclusion body product.

(d) **Purificati n of SCR 1 + 2 + 3 from the solubilised inclusi n body**

A column (i.d., 16 mm; h, 10 mm) of S-Sepharose Fast Flow was prepared and connected into an FPLC (Pharmacia) system. The column was equilibrated with 20 mM Tris/8M urea/1 mM EDTA/50 mM 2-mercaptoethanol pH 8.5. 10 ml of

thawed, solubilised inclusion body, prepared as described in Example 2c, was applied to the column and washed through with equilibration buffer. The column was then developed with a linear gradient to 1M NaCl (in equilibration buffer) followed by rinses with 1M NaCl and 2M NaCl (also in equilibration buffer). All the chromatography was at 1.0 ml min $^{-1}$ and at room temperature.

Analysis by SDS PAGE/protein staining of the fractions collected during the chromatography indicated that virtually all the SCR 1 + 2 + 3 polypeptide had absorbed to the column and had been dissociated by the 1M NaCl - containing buffer. The appropriate fractions were stored at -40°C.

Subsequent assay for protein content of the peak fraction using the Bradford protein assay and a bovine serum albumin standard showed it contained 2.8 mg protein.

(e) **Folding**

S-Sepharose - purified SCR 1 + 2 + 3 that had been purified in a similar way to that described in Example 2d and stored at -40°C was thawed and 0.4 ml buffer-exchanged into 0.05 M formic acid using Sephaex G25 (P10). The absorbance at 280 nm of the buffer-exchanged solution was determined as 0.52, and, using ϵ = 34000 and appropriate correction factors for dilution, the protein concentration of the original preparation (prior to buffer-exchange) was calculated to be 0.6 mg/ml.

Based on this figure, 1.7 ml S-Sepharose - purified protein was diluted with 0.85 ml 20 mM Tris/8M urea/50 mM 2-mercaptoethanol/1M NaCl pH 8.5 to give a 0.4 mg/ml solution, on which the folding was carried out.

Folding was effected by a series of dilutions, using cold diluent at all times.

At t=0 h, 0.8 ml SCR1 + 2 + 3 (0.4 mg/ml) was added to 0.8 ml 20 mM Tris/1M urea/5 mM EDTA/3 mM 2-mercaptoethanol pH 8.0 ('diluent') in a 30 ml polystyrene universal container. The solution was mixed thoroughly by gentle swirling and left static, capped, in a cold room (approx. 2 to 3°C).

At 1 h, 1.6 ml diluent was added and mixed.

At 2 h, 3.2 ml diluent was added and mixed.

At 4 h, 6.4 ml diluent was added and mixed.

The solution was left a further 20 h in the cold room, then ultrafiltered (YM5, Amicon Ltd) to approx. 1.4 ml. This was buffer-exchanged into 0.1M NH_4HCO_3 (2.5 ml) using Sephadex G25 (PD 10) in the cold room. The eluate was aliquoted and was stored at -40°C or lyophilised.

The product containing SCR 1 + 2 + 3 was analysed by SDS PAGE, followed by protein staining. In both non-reduced and reduced (with 2-mercaptoethanol 5% (v/v)) samples there was a single major band. The molecular weight of the reduced band, compared to reduced protein standards of known M_r, was approx. 24,000. The

non-reduced protein (band) had a slightly faster mobility than the reduced protein (band).

The product was analysed in a functional haemolytic assay utilising antibody-sensitized sheep erythrocytes (Method (xvi)). The product showed concentration-dependent inhibition of the complement - mediated lysis of the erythrocytes with an IH50 around 0.4 µg/ml.

(f) **Folding**

Preparation, folding, processing and analysis were carried out exactly as described in Example 2e except

(1) the diluent for the folding was 20 mM ethanolamine (pH 10.0)

(2) the folded solution was ultrafiltered to a final volume of 1.55 ml, and

(3) the IH50 figure was determined as about 0.6 µg/ml.

(4) the recovery of product was approx. 100 per cent.

(g) **Determination of N-terminal sequence of SCR 1 + 2 + 3**

1ml samples of growing *E.coli* BL21 (DE3) containing plasmid pDB1013-5-4 were removed 3 hours post-induction with 1 mM IPTG as described in Example 2b. These samples were spun in an eppendorf centrifuge and the resultant pellets each resuspended in 200µl of reducing buffer (100 mM Tris pH6.8/200 mM dithiothreitol/4% (^{w}/v) SDS/2% (^{w}/v) bromophenol blue and 20% (^{v}/v) glycerol and boiled for 5 minutes. 25µl samples were applied to a 14% polyacrylamide gel. When the electrophoresis was complete the proteins were transferred to a ProBlott membrane (Applied Biosystems) using a Sartoblot electroblotting apparatus (Sartorius) at 0.8 mA/cm^2 for 1 hour 40 mins using CAPS (3-[cyclohexylamino]-1-propanesulphonic acid) transfer buffer. After transfer the ProBlott membrane was stained (0.1% (^{w}/v) Coomassie Blue R-250/40% (^{v}/v) methanol/1% ^{v}/v acetic acid) for 20 seconds and destained using 50% (^{v}/v) methanol. A band corresponding to a M_r approx 23,000 protein was excised and the N-terminal sequence determined using a Blott cartridge in an Applied Biosystems 477A Protein Sequencer.

The sequence of the first 20 amino acids was found to agree with the predicted sequence except that residue 3 could not be identified by the sequencing protocol used.

Example 3 Expression and purification of SCR 1 + 2 + 3 from a fermentation vessel

(a) **Fermentati n of *E. coli* harbouring the plasmid pDB1013-5-4**

E coli BL21 (DE3) : pDB1013-5-4 was recovered from storage in liquid nitrogen by thawing a vial containing 1 ml of the culture and this was used to inoculate 100 ml of seed medium (NCYZM) containing ampicillin at 75 µg/ml. The

primary and secondary seed stage fermentations were carried out in plain 500 ml shake flasks batched with 100 ml aliquots of NCYZM medium. The primary and secondary seed fermentation conditions were as follows : 37°C, 230 rpm on an orbital shaker with a 50 mm throw. The primary seed incubation time was 2 hours. The primary seed culture was used to inoculate secondary seed fermentation medium at 0.1% (v/v). The secondary seed was incubated for 14.5 hours.

Two 15 litre Biolafitte fermenters were each batched with 10 litres of NCYZM medium and 0.01% (v/v) Dow Corning DC1510 antifoam. The vessels plus media were sterilised using steam to 121°C for 45 minutes. Ampicillin sterilised by microfiltration (0.2 µm) was added aseptically to the vessel media to a final concentration of 150 µg/ml. The fermenters were inoculated at a level of 3% (v/v) from pooled secondary seed culture. The final stage incubation conditions were 37°C, agitator 400 rpm, airflow 5 l/min (0.5 vvm). The final stage fermentations were sampled aseptically pre-inoculation, at 0 hours and thence every half hour. The samples were monitored for increases in optical density (600 nm). When the OD600 was ≥ 0.5, IPTG was added to give a final concentration of 1 mM. The fermentations were incubated for a further 3 hours.

The cells were recovered by centrifugation using 7000 g for 25 minutes. The total cell yield (wet weight) was 49.8 grammes.

(b) **Inclusion body isolation.**

Inclusion bodies from 23 g (wet weight) cell pellet were isolated and solubilised essentially as described in Example 2.

(c) **Purification of denatured SCR 1 + 2 + 3**

The volume of solubilised inclusion body from Example 3b was approx. 800 ml. To this viscous solution was added SP-Sepharose FF (100 ml gel bed, water washed and suction dried). The mixture was swirled vigorously and left static for 1h at room temperature. The supernatant was decanted and stored at -40°C. The remaining slurry was resuspended to a uniform suspension and poured into a glass jacket (i.d., 41.5 mm) and allowed to settle into a packed bed. This packed bed was connected into a low pressure chromatography system at 4°C and equilibrated with 0.02M Tris/8M urea/0.05M 2-mercaptoethanol/0.001 M EDTA pH 8.5. When the A_{280} of the eluate had minimised, the buffer was changed (step-wise) to the equilibration buffer additionally containing 1 M NaCl. A single A_{280} peak was eluted, in a volume of 90 ml (equivalent to approx. 1 Vt). The solution was clear and colourless and was estimated, by A_{280} determination of a buffer-exchanged sample (using an $\varepsilon = 25{,}000$), to contain about 300 mg target protein. By SDS PAGE followed by protein stain the target protein was the major band present. The 90 ml product was stored at -40°C.

(d) **F lding and further purificati n.**

18 ml of the above product (nominal 60 mg) was diluted with 12 ml 0.02 M Tris/8M urea/1 M NaCl/0.05M 2-mercaptoethanol pH 8.5. The product (30 ml) was added as 5 ml aliquots at 1 min intervals to 930 ml freshly prepared, cold 0.02 M ethanolamine/0.001 M EDTA, with swirling, and left static for 1 h/4°C. Then reduced glutathione was added to 1 mM (by addition of 9.6 ml 0.1M GSH) and oxidised glutathione was added to 0.5 mM (by addition of 9.6 ml 0.05M GSSG). The solution was clear and was left static in the cold for approx. 70 h. The solution was then ultrafiltered using a YM10 membrane to a final retentate volume of about 10 ml; this retentate was cloudy. It was mixed with 90 ml 0.1 M NaH_2PO_4/1 M $(NH_4)_2SO_4$ pH 7.0 (Buffer A) at room temperature and then centrifuged at 4000 rpm for 20 min. The supernatant was decanted and SCR 1 + 2 + 3 protein isolated by chromatography of the supernatant on Butyl Toyopearl 650 S.

The column of Butyl Toyopearl (Vt ~ 12 ml) was equilibrated with Buffer A. The 100 ml supernatant was applied to the column and the column washed with Buffer A. It was then developed with a linear gradient of 100% Buffer A to 100% 0.1 M NaH_2PO_4 pH 7.0. All the chromatography was at room temperature at approx. 30 cmh^{-1}.

A major A_{280} peak was eluted during the gradient. Fractions spanning the peak were analysed by SDS PAGE followed by protein stain. The most concentrated fractions of the peak contained essentially pure SCR 1 + 2 + 3 and were active in the haemolytic assay (Method (xvi)) (IH_{50} ~ 0.3 µg/ml). They were stored at -40°C.

Example 4 Formulation of Butyl Toyopearl purified SCR 1 + 2 + 3.

Batches of SCR 1 + 2 + 3 that had been expressed, folded and purified in similar ways to batches described in Examples 2 and 3 and further purified by ammonium sulphate treatment and Butyl Toyopearl chromatography essentially as described in Example 3d were formulated into a useable product as follows.

Three such Butyl Toyopearl products were pooled to give a volume of about 31 ml. All 31 ml were buffer-exchanged into 0.05 M formic acid (prepared using 0.2 µm-filtered 'MilliQ' water) using a column of Sephadex G25. All the chromatography was at 50 cmh^{-1} at 4°C. The eluate from the column was monitored at 280 nm and the Vo fraction was collected as a single fraction. The bulk of this fraction was lyophilised in aliquots.

Analysis of the Vo pool prior to lyophilisation by both SDS PAGE/stain and C8 reverse phase HPLC showed it to be essentially pure target protein. The pool demonstrated anti-haemolytic activity (IH_{50} approx. 0.3 µg/ml) and the endotoxin content was low (< 1ng/mg).

One of the lyophilised aliquots was shown to be soluble at 10 mg ml^{-1} in phosphate-buffered saline and showed complement inhibitory activity in the haemolytic assay (Method xvi); the IH50 was 0.3µg/ml.

Another of the lyophilised aliquots was examined to determine the disulphide bridge pattern. All six correct (as predicted on the basis of a consensus SCR motif) disulphides were detected.

Example 5 Effect of SCR 1+2+3 on IgG-mediated activation of the classical pathway of complement, as measured by C3a release

Inhibition of heat aggregated IgG activated serum was carried out as described in Method (xvii). Heat aggregated IgG activates the classical pathway of complement. Different concentrations (typically 4 - 1000 µg/ml) of inhibitor were incubated with serum in the presence of heat aggregated IgG and the % inhibition of activation at each concentration was determined. The IC50 of SCR 1+2+3 was determined as approximately 22 µg/ml indicating that SCR 1+2+3 can inhibit the classical pathway of complement.

Example 6 Effect of SCR 1+2+3 on zymosan A-mediated activation of the alternative pathway of complement, as measured by following C3a release.

Inhibition of zymosan A activated serum was carried out as described in Method (xvii). Different concentrations of SCR 1+2+3 (typically in the range 1 - 1000 µg/ml) were incubated with serum in the presence of zymosan A and the % inhibition of activation at each concentration was determined. From several different experiments the IC50 was determined as 20 - 40 µg/ml indicating that SCR 1+2+3 can inhibit the alternative pathway of complement.

Example 7 Effect of SCR 1+2+3 on zymosan A-mediated activation of the alternative pathway of complement, as measured by C5a release.

Inhibition of zymosan A activated serum was carried out as described in Method (xvii) and assayed as described in Method (xviii). Different concentrations of SCR 1+2+3 (typically in the range 4 - 700 µg/ml) were incubated with serum in the presence of zymosan A and the % inhibition of activation at each concentration was determined. From several different experiments the IC50 was determined as approximately 20 - 30 µg/ml, indicating that SCR 1 + 2 + 3 can inhibit the alternative pathway of complement.

Example 8 End t xin content determination of purified, folded and formulated SCR 1 + 2 + 3

A batch of final product SCR 1 + 2 + 3 was prepared essentially as described in Example 4 above and was measured for endotoxin content using a method based on the gel-clot reaction of limulus amoebocyte lysate (LAL) (Atlas Bioscan Ltd.). The sensitivity of the assay was 0.125 EU/ml and this was checked by titration against a doubling dilution series prepared from standard *E. coli* endotoxin supplied with the LAL kit.

10-fold dilutions of ~ 1.3 mg/ml SCR 1 + 2 + 3 protein stock were tested in quadruplicate for their effect on LAL by adding 10µl of sample to 10µl LAL. After 1h at 37°C the mixtures were tested for either clotting or remaining liquid. (Solutions that contain at least 0.125 EU endotoxin will clot this LAL preparation.) After taking into account the results of simultaneous tests designed to test for interference, it was concluded that the endotoxin content of the SCR 1 + 2 + 3 protein preparation was < 12.5 EU/ml, equivalent to approx. < 1 ng/mg protein.

Example 9 Expression, folding, purification, and formulation of MR122 -> K196 of CR-1 (SCR 3)

General Points

The sequence for SCR 3 corresponding to amino acid 122 and ending at amino acid 196 of mature human complement receptor 1 was designed such that the 5' end of the gene contained an *Nde*I restriction endonuclease site. This site comprises an ATG start codon to give the initiating methionine required for the start of mRNA translation and allows the placement of the gene an optimum distance from the Shine-Dalgarno ribosome binding site of pBROC413. This codon is followed immediately by the gene coding for SCR 3 starting with arginine 122 of mature human complement receptor 1. The 3' end of the gene finishes with a codon for lysine 196 followed by two stop codons followed by a *Hin*dIII site.

The DNA coding for SCR3 was modified for optimum codon usage in *E. coli* as described in the methods. The gene was also altered to incorporate unique restriction endonuclease sites. This was carried out in the following way. Restriction endonucleases that do not cut pBROC413 and were commercially available were identified. The DNA sequence of these restriction endonuclease sites was then translated into all three reading frames and the codon usage examined. Sites that contained codons that are rarely used by *E. coli* were discarded. The remaining sites were examined for their translated sequence and if that sequence matched with SCR 3, the restriction site was incorporated into the sequence.

(a) **Constructi n of plasmid pBROC435 encoding SCR 3**

The construction of pBROC435 is described in Example 2a

(b) **Expression f SCR 3 from pBROC435**

pBROC435 was transformed by electroporation into *E.coli* BL21(DE3) and resulting colonies analysed by restriction digestion of mini-plasmid DNA preparations. Single colonies were inoculated into universals containing 10 ml of L broth or NCYZM medium and 50 - 75 µg/ml ampicillin and allowed to grow overnight at 37°C, 220 r.p.m. Typically 4ml of overnight cultures were used to inoculate each of 2 L conical flasks containing 500 ml of NCYZM medium, 150 µg/ml ampicillin; cultures were grown at 37°C, 230r.p.m. until A_{600} was 0.5 absorbance units. Cultures were induced with 1 mM IPTG and allowed to grow a further 3 hours under the same conditions. The cultures were centrifuged (approx, 8000g/10 min) and the supernatants discarded. The cell pellets were stored at -40°C.

(c) **Isolation of solubilised inclusion bodies**

The frozen cell pellet of *E. coli* (from 3 l growth culture in NCYZM) described in Example 9b was allowed to thaw at room temperature for 2 h and was then resuspended in 50 mM Tris/50mM NaCl/1mM EDTA/0.1 mM PMSF pH 8.0 (90ml). The suspension was transferred to a 200 ml glass beaker and sonicated (Heat Systems - Ultrasonics W380; 70 Watts, 50 x 50% pulse, pulse time = 5 sec.). The sonicate was immediately centrifuged (6000g/4°C/10 min) and the supernatant was discarded. The pellet, containing the inclusion bodies, was resuspended in 20 mM Tris/8 M urea/50 mM 2-mercaptoethanol/1 mM EDTA/0.1mM PMSF pH 8.5 (300ml) with gentle pipetting to mix. After mixing, the solution was left static at room temperature (approx. 23° C) for 1 h. The resulting solution was centrifuged (2000 g at 4° C for 10 min) to remove material that failed to solubilise. The supernatant of this spin was retained at -40° C as the solubilised inclusion body product.

(d) **Purification of SCR3 from the solubilised inclusion body**

A column (i.d., 32 mm; h, 32mm) of Q-Sepharose Fast Flow (Pharmacia) was prepared and equilibrated with 20 mM Tris/8M urea/50 mM 2-mercaptoethanol pH 9.0. 200 ml of thawed, solubilised inclusion body, prepared as in Example 9c, was applied to the column and washed through with equilibration buffer. The column was connected to an FPLC system and developed via a stepwise gradient of 0.1, 1.0, 2.0M NaCl (also in equilibriation buffer). All chromatography was at 2.0 ml min^{-1} and at room temperature.

Analysis by SDS PAGE/protein staining of the fractions collected during the chromatography indicated that virtually all the SCR3 did not bind to the column. Many other proteins had absorbed to the column however and had been dissociated by the 0.1M and 1M NaCl - containing buffers. The purity of SCR3 in the

unadsorbed fraction was estimated to be about 80%.

(e) **Folding of SCR3**

Q-Sepharose-purified SCR3 that had been purified as described in Example 9d and stored at -40°C was thawed and was folded by a series of dilutions, using cold diluent. At t=0, 100 ml 20 mM Tris/1M urea/5mM EDTA/3mM 2-mercaptoethanol pH 8.0 (diluent) were added to 100ml SCR3. At this stage the solution was turbid in appearance. The solution was mixed thoroughly by gentle swirling and left static, capped, in a cold room (2-3°C). At 1 h, 200ml diluent was added and mixed, final volume = 400 ml. At 2 h, 400ml diluent was added and mixed, final volume = 800ml. At 4 h 800ml of diluent was added and mixed, final volume = 1.6L. The solution was left for a further 20 h in the cold room. The solution now appeared clear, and it was stored at -40°C in aliquots.

(f) **Formulation of SCR3**

50 ml of SCR3 prepared as in Example 9e were thawed and ultrafiltered to 3.5 ml using a 2000 Da cut-off membrane (Amicon). 2.5 ml of the concentrate was buffer-exchanged into 0.1M NH_4HCO_3 (3.0 ml) using Sephadex G25 (PD 10). Subsequent analysis for protein content using the molar extinction coefficient of 11000 showed this sample contained approx 0.24mg/ml.

Analysis of this material by SDS PAGE/protein staining indicated that the protein was about 80% pure. Samples reduced with 2-mercaptoethanol had a lower electrophoretic mobility suggesting the presence of disulphide bonds in SCR3.

Analysis of this sample in the haemolytic assay (Method (xvi)) showed it had an IH50 of approx. 10 - 20 μg/ml.

(g) **Determination of N-terminal sequence of expressed SCR3**

200μl SCR3 prepared and formulated in 0.1M NH_4HCO_3 as in Example 9f was precipitated with 800μl cold acetone in a cardice/ethanol bath for 60 mins. The sample was then spun in an Eppendorf centrifuge (approx 10,000g/20 mins) and the resultant pellet resuspended with heating in sample buffer containing 5% (v/v) 2-mercaptoethanol. 30 μl samples were electrophoresed on a 4 to 20% SDS-containing polyacrylamide gradient gel. When the electrophoresis was complete the proteins were transferred to a ProBlott membrane (Applied Biosystems) using an electroblotting apparatus at 200mA for 2 h using CAPS in 10% methanol/90%H_2O (v/v)) transfer buffer. After transfer the ProBlott membrane was stained (0.1%(w/v) Coomassie Blue), destained, rinsed and air dried according to the manufacturer's instructions. Sections of the membrane were excised and used for N-terminal sequencing.

The sequence of the first 20 amino acids of the major band was as expected for SCR3 with the exception of residue 5, which could not be identified.

(h) **Preparati n, f lding and f rmulati n f SCR3**

Preparation and folding were carried out exactly as described in Example 9a-9e. 400 ml of folded SCR3 was ultrafiltered through a 30 KDa cut-off filter (Amicon) at 4° C. Samples of the ultrafiltrate were processed in two ways.

1. 50 ml were ultrafiltered using a 2 KDa cut-off membrane to a final volume of 3.5ml and buffer-exchanged into 0.05 M formic acid (6.7 ml) using Sephadex G25 (PD10) columns. The total amount of SCR3 estimated by the absorbance at 280 nm was 0.6mg. Analysis by SDS PAGE/protein staining indicated that the protein had a purity of about 95%. The sample was freeze-dried and stored at -40°C.

2. 100ml of the ultrafiltrate were adjusted to pH 5.5 with HCl. The sample was applied to a Mono S column (1ml) at 1.5 ml min^{-1} and washed through with equilibration buffer (20mM Tris.HCl pH 5.5). The column was then developed with a step gradient of 0.1, 1.0 and 2.0M NaCl (also in equilibration buffer). All remaining chromatography was at 1.0 ml min^{-1} and at room temperature.

Analysis by SDS PAGE/protein staining of the fractions collected during the chromatography demonstrated that the major band dissociated at 1M NaCl contained SCR3 at about 95% purity.

Example 10 Expression, folding, purification and formulation of MR122-S253 of CR-1 (SCR 3 + 4)

(a) **Construction of plasmid pDB1019 encoding SCR 3 + 4**

The DNA coding for SCR 3 + 4 was constructed from the plasmids pBROC435 (Example 2) and pDB1018-1 (Example 11) which carry the genes coding for SCR 3 and SCR 1 + 2 + 3 +4 respectively. The SCR 4 coding unit was excised from pDB1018-1 and ligated onto the end of the SCR 3 coding unit in pBROC435.

pDB1018-1 was digested with *Spe*I and *Hin*dIII and separated on a 1 % agarose gel. The band which codes for SCR 4 (~ 245 bp) was excised from the gel and purified using the QIAEX extraction kit. Plasmid pBROC435 was also cut with *Spe*I and *Hin*dIII, separated on 1 % agarose, excised from the agarose and purified with the QIAEX kit. The SCR 4 coding DNA was then ligated into the cut pBROC435 plasmid to give pDB1019. This DNA was used to transform *E. coli* HB101 made competent with $CaCl_2$. Transformants were analysed by restriction analysis using *Eco*RI and *Hin*dIII. Clones carrying the correct sized insert were used for expression studies.

(b) **Expression of SCR 3 + 4 from pDB1019-1C**

pDB1019 was transformed into *E. coli* BL21(DE3) made competent with $CaCl_2$ and the resulting colonies were analysed by restriction digestion of mini-

plasmid DNA preparations. Plasmid pDB1019-1C was identified as carrying the correct sized insert. Single colonies of *E. coli* BL21(DE3) carrying pDB1019-1C were inoculated into ten universals containing 10 mls of NCYZM medium and 75 μg/ml ampicillin and allowed to grow overnight at 37°C, 240 r.p.m. The overnight cultures were then used to inoculate eight 2 L conical flasks (5 ml/flask) containing 500 ml of NCYZM medium, 150 μg/ml ampicillin. Cultures were grown at 37°C, 240 r.p.m. until A_{600} was 0.5 absorbance units. At this point cultures were induced with 1 mM IPTG and allowed to grow a further 3 hours under the same conditions. The cultures were centrifuged (approx. 8000g/ 10 mins) and the supernatants were discarded. The cell pellets were stored at -40°C.

(c) Isolation, purification, folding and formulation of SCR 3 + 4

The methods used generally follow those described earlier for the preparation of SCR 1 + 2 + 3.

Isolation of solubilised inclusion bodies from cell pellet derived from 2l culture was carried out as described in Example 2c. The volume of solubilisate was 200 ml.

Some of the contaminating (host) *E. coli* proteins were removed from the preparation by adsorption onto S-Sepharose, either in a batch process or by column chromatography, using systems similar to those described in Example 2d. The protein present in the unadsorbed fractions was shown by SDS PAGE/stain to contain significant amounts of SCR 3 + 4 protein. About half of these fractions were ultrafiltered using a YM1 (Amicon) membrane to approx. 35 to 40 ml. This retentate was estimated to contain about 0.3 mg protein/ml (based on A_{280} determination of a buffer-exchanged sample, using ε= 21,000). 10.5 ml of the retentate was mixed with 325 ml cold 20 mM ethanolamine and left static at 4°C for 1 hour. Then reduced glutathione was added to 1 mM (by addition of 3.4 ml 100mM GSH) and oxidised glutathione was added to 0.5mM (by addition of 3.4ml 50mM GSSG). The solution was mixed and left static at 4°C for ~ 72 h. The solution was clear. The solution was then ultrafiltered using a YM1 membrane to a retentate of 5 ml. The retentate was divided in two and buffer-exchanged into either 20 mM ethanolamine or 50 mM formic acid using Sephadex G25 (PD10 columns).

Analysis of the formic acid SCR 3 + 4 product by reverse phase HPLC and by SDS PAGE followed by protein staining showed only one major protein species (> 90% pure). The protein concentration was estimated to be 0.3 mg/ml using A_{280} determinations. The product was active in the haemolytic assay (Method (xvi)); the IH50 value was approx. 30 μg/ml

Example 11 Construction, expression, folding, purification and formulati n of MQ1-S253 of CR-1 (SCR 1 + 2 + 3 + 4)

General p ints

Two constructs were prepared by making a plasmid encoding SCR 1+2, incorporating SCR3 and finally adding SCR4. The two constructs encoded consensus SCR1 to 4 and the R235H mutation of SCR1 to 4 (Example 12).

A plasmid containing the SCR 1 + 2 + 3 + 4 coding unit was constructed by adding the DNA encoding SCR 4 onto the construct coding for SCR 1 + 2 + 3 (Example 2). For convenience of DNA manipulation, the SCR 4 DNA coding unit was made by synthesising the DNA encoding the last 17 amino acids of SCR 3 followed by the DNA coding for the linker region followed by SCR 4. This DNA started at the *Spe*I site of the SCR 1 + 2 + 3 coding construct which corresponds to T175 of mature CR-1 followed by the DNA coding for the linker region followed by SCR 4 ending on the codon corresponding to S253 followed by two stop codons and a *Hind*III site. As for the previous constructs the DNA encoding SCR 4 was altered for optimised codon usage and restriction sites as previously described in Example 1. This unit of DNA was ligated to the plasmid coding for SCR 1 + 2 + 3 which had been cut with *Spe*I and *Hind*III to give a construct coding for SCR 1 + 2 + 3 + 4.

(a) **Construction of plasmid pDB1018 encoding SCR 1 + 2 + 3 + 4**

Oligonucleotides (Table 1; oligos 21 - 26 coding for SCR4) were synthesised as 3 complementary pairs of 68-90 mers that could be ligated in a unique fashion via complementary 8 bp overhangs between the pairs of oligonucleotides. The 3 complementary pairs of oligonucleotides were designated Pair E (oligos 21, 22), Pair F (oligos 23, 24) and Pair G (oligos 25, 26). Pair E which corresponds to the 5' end of the gene contained a *Spe*I restriction site overhang and Pair G contained a *Hin*d III restriction site overhang at the 3' end. All oligonucleotides apart from 22 and 24 were purified by electrophoresis through a denaturing polyacrylamide gel followed by reverse phase chromatography (C_{18}). Oligonucleotides 22, 23, 24 and 25 were kinased before annealing to their complementary oligonucleotides. The oligonucleotides were ligated pair E to pair F to pair G to form the gene coding for part of SCR3 and the whole of SCR4 which for convenience will be called the SCR4 gene in the subsequent text.

The DNA coding for SCR4 was initially amplified by PCR using two oligonucleotides (Table 1; oligos 17 and 18) complementary to the two strands of DNA. Both oligonucleotides contained 5' unmatched ends that contained 6bp of random sequence followed by the sequence of either *Spe*I (oligo 17) or *Hind* III (oligo 18) restriction sites followed by 18 bp complementary to the gene. Following PCR a band of approximately 250 bp was visualised on horizontal agarose gel

electrophoresis, which was excised and purified on DEAE NA45 membranes. This DNA was used for a second PCR amplification using nested primers that had been moved inwards by four nucleotides at their 5' ends (Table 1; oligo 19, oligo 20). These oligo's incorporated the *Spe*I and *Hind*III restriction sites but now only had 2 nucleotides beyond the end of each restriction site. Following PCR a band of approx. 250 bp was visualised on horizontal agarose gel electrophoresis. This band was excised and purified using the QIAEX agarose gel extraction kit.

The DNA for SCR 4 was blunt-end ligated to itself following kinasing. The multimers formed were visualized by horizontal agarose gel electrophoresis and the bands excised and purified using the QIAEX agarose gel extraction kit. The DNA was then cut with *Spe* I and *Hind* III and ligated into pDB1013-5-4 that had been cut with the same enzymes to produce pDB1018 (Fig.3). The vector was transformed into *E.coli* HB101 made competent with calcium chloride. Mini-plasmid preparations were made and plasmid DNA analysed by digestion with *Nde* I, *Hind* III, *Stu* I, *Spe* I and *Kpn* I. The plasmids with the correct restriction maps were analysed by DNA sequencing of both strands across the gene encoding SCR4. Two plasmids were selected for further study. pDB1018-1, which encoded MQ1-S253 (consensus SCR1 to 4) and pDB1018-6, which encoded the R235H mutant of MQ1-S253. The amino acid sequences of the two polypeptides encoded by pDB1018-1 and pDB1018-6 are shown in Table 2.

Taking the first residue as being the A of the ATG initiating codon, DNA sequencing revealed that residue 600 of pDB1018-6 had been altered from G -> A. This is a silent mutation and does not alter the amino acid at this position.

(b) **Expression of MQ1-S253 from pDB1018-1**

pDB1018-1, constructed as described in Example 11a, was transformed into calcium chloride competent *E.coli* BL21(DE3). Single colonies were inoculated into universals containing 10ml of NZCYM medium and 75 μg/ml ampicillin and allowed to grow overnight at 37°C, 230 r.p.m. 3ml of overnight culture were used to inoculate each of 8 x 2 litre conical flasks containing 500ml of NZCYM medium, 150 μg/ml ampicillin; cultures were grown at 37°C, 230 r.p.m. until A_{600} reached 0.5 absorbance units. The cultures were induced with 1mM 1PTG and allowed to grow for a further 3 hours under the same conditions. The cultures were centrifuged (approx. 7000g/10 mins/4°C) and the supernatants discarded. The cell pellets were stored at -40°C.

(c) **Isolation f solubilised inclusion bodies**

The frozen cell pellets of *E.coli* BL21(DE3) (pDB1018-1) each equivalent to 1 litre of culture prepared as described in Example 11b were allowed to thaw at 0-4°C over 2 hours. The pellets were resuspended in 50mM Tris/50mM NaCl/1mM

EDTA/0.1mM PMSF pH 8.0; 30ml for each litre pellet. Each suspension was transferred to a 100ml glass beaker and sonicated (Heat systems - Ultrasonics W380; 70 Watts, 50 x 50% pulse, pulse time = 5 seconds). The sonicates were pooled and immediately centrifuged (6,000 g/ 4°C/10 mins) and the supernatant discarded. The pellet containing the inclusion bodies was resuspended in 20 mM Tris/8 M urea/50mM 2-mercaptoethanol/1mM EDTA/0.1mM PMSF pH 8.5 (400 ml), thoroughly mixed and left static at room temperature (approx. 23°C) for 1 hour.

(d) **Purification of MQ1-S253 from the solubilised inclusion body.**

30ml of S-Sepharose FF that had been washed with deionised water and suction dried was added to the inclusion body solution described in Example 11c, and vigorously shaken for 30 seconds. The S-Sepharose mixture was left static at room temperature (23°C) for 1.5 hours and then the supernatant was discarded. The remaining slurry was packed into a column (id, 4.1cm). The column was equilibrated using 20mM Tris/8M urea/50mM 2-mercaptoethanol/1mM EDTA/0.1mM PMSF pH 8.5 at 60 cmh^{-1}, 4°C. MQ1-S253 protein was eluted using the equilibration buffer containing 1M NaCl. Analysis by SDS PAGE/protein staining of the fractions collected during the chromatography indicated that virtually all the target protein had adsorbed to the column and had been dissociated by the 1M NaCl wash. The appropriate fraction was stored at -40°C.

(e) **Folding and formulation**

Based on a molar extinction coefficient of 25,000 and A_{280} values determined in 50mM formic acid, 60 mg of the S-Sepharose purified unfolded protein described in Example 11d was folded and formulated as follows :-

8.0ml of solution (equivalent to 60 mg protein) was diluted with 22ml cold 20mM Tris/8M urea/50mM 2-mercaptoethanol/1M NaCl/1mM EDTA/0.1mM PMSF pH8.5, to give 30ml of a 2.0 mg/ml solution. The 30ml was diluted rapidly with constant stirring into 930ml cold (0-4°C) freshly prepared 20mM ethanolamine. The solution was left static at 0-4°C for 1 hour. Reduced glutathione was added to 1mM (by addition of 9.6ml of 100mM stock) and then oxidised glutathione was added to 0.5mM (by addition of 9.6ml of 50mM stock). The solution was left static at 0-4°C for a further 48 hours and then ultrafiltered using a stirred cell (Amicon) and a YM10 membrane (Amicon, nominal 10,000 Da molecular weight cut-off) to approx. 29ml. The ultrafiltered retentate was buffer exchanged into 50 mM formic acid using Sephadex G25 (i.d., 26mm; h, 245mm Vt, 123ml) and a flow rate of 50 cmh^{-1} to a final volume of 40ml. Using a molar extinction coefficient of 25,000 for the protein 51mg of protein was recovered. The purified protein gave an IH_{50} value (see Method xvi) of approximately 2 µg/ml.

(f) **Further purification and formulation of SCR1 + 2 + 3 + 4.**

Folded SCR1 + 2 + 3 + 4 (nominal 25mg) in 50mM formic acid prepared essentially as described in Example 11e was lyophilised. The lyophilisate was resolubilised in 20mM ethanolamine (10ml) to give a cloudy solution. The 10ml were then added to 90ml 0.1M NaH_2PO4/1M $(NH4)_2SO4$ pH 7.0, thoroughly mixed, and then clarified by centrifugation (4000 rpm/20 min). The supernatant (100ml) was decanted and was chromatographed on Butyl Toyopearl (exactly as described for SCR1 + 2 + 3 in Example 3d). The peak A_{280} fractions, eluting at about 100% of the 1M NaCl-containing buffer, were pooled and buffer-exchanged using Sephadex G25 into 50mM formic acid. The Vo pool (29.5ml) was lyophilised in aliquots.

The purity of the protein was assessed by SDS PAGE followed by protein staining and by C8 reverse-phase HPLC; the protein was estimated to be >95% pure. One of the lyophilised aliquots was resolubilised to 4mg protein/ml in 0.1M Hepes/0.15M NaCl pH7.4. The product showed activity in the haemolytic assay (Method (xvi)); the IH50 was calculated to be 0.3 µg/ml.

Another of the lyophilised aliquots was examined to determine the disulphide bridge pattern using proteolytic digestion and peptide identification by amino acid sequencing. All eight correct (as predicted on the basis of a consensus SCR motif) disulphides were detected.

Example 12 Expression, isolation, folding and formulation of purified MQ1-S253 (R235H)

(a) **Expression of MQ1-S253 (R235H)**

pDB1018-6 (prepared as described in Example 11a) was transformed into calcium chloride competent *E. coli* BL21(DE3). Single colonies were inoculated into universals containing 10 mls of NCYZM medium and 50 µg/ml ampicillin and allowed to grow overnight at 37°C, 220 r.p.m. The overnight cultures (approx. 3ml) were used to inoculate each of 2 l conical flasks containing 500 ml of NCYZM medium, 150 µg/ml ampicillin; cultures were grown at 37°C, 220 r.p.m. until A_{600} was 0.5 absorbance units. Cultures were induced with 1 mM IPTG and allowed to grow a further 3 hours under the same conditions. The cultures were centrifuged (approx. 8000g/10 min/4°C) and the supernatants discarded. The cell pellets were stored at -40°C.

(b) **Isolation of solubilised inclusion bodies and purification of unfolded MQ1-S253 (R235H)**

Frozen cell pellet of *E. coli* BL21 DE3 (pDB1018-6) (2 litre culture) described in Example 12a was allowed to thaw at 4°C for 2 h and was then resuspended in 50 mM Tris/50 mM NaCl/1 mM EDTA/0.1 mM PMSF pH 8.0 (66

ml). The suspension was transferred to a 250 ml glass beaker and sonicated (Heat Systems - Ultrasonics W380; 70 Watts, 30 x 50% pulse time = 5 seconds). The sonicate was immediately centrifuged (6000g/4°C/10 min) and the supernatant was discarded. The pellet, containing the inclusion bodies, was resuspended by vigorous swirling in 20 mM Tris/8 M urea/50 mM 2-mercaptoethanol/1 mM EDTA/0.1 mM PMSF pH 8.5 (200 ml) and left static at room temperature (approx. 23°C) for 1.5h. Water-washed, suction-dried S-Sepharose (equivalent to approx. 25 ml packed bed volume) was added to the 200 ml solubilised inclusion body and the mixture swirled vigorously to wet the Sepharose beads thoroughly. The mixture was left static at room temperature for 1h. The supernatant (approximately 150 ml) was decanted and discarded. The slurry remaining was resuspended to a uniform suspension by swirling and then poured into a 32 mm (i.d.) glass jacket and allowed to settle. The gel bed was connected into a low pressure chromatography system and was equilibrated with 20 mM Tris/8 M urea/1mM EDTA/50mM 2-mercaptoethanol pH 8.5 at 4°C until the A_{280} baseline stabilised. The column was then developed with equilibration buffer containing 1M NaCl. All the chromatography was at approx. 1 ml min^{-1}. Analysis by SDS PAGE/protein staining of the fractions collected during the chromatography indicated that most of the MQ1-S253 (R235H) polypeptide had adsorbed to the column and had been dissociated by the 1M NaCl - containing buffer wash and that the purity of the material was about 90%.

A sample of the pool was buffer-exchanged into 50mM formic acid using Sephadex G25 column to allow some assays to be carried out.

Amino acid analysis of the pool of the MQ1-S253 (R235H) - containing fractions gave a total protein content of about 120 mg.

(c) **Folding and formulation of SCR 1 + 2 + 3 (R235H)**

Based on A_{280} values and a molar extinction coefficient of 25,000 for the protein in 50mM formic acid, 20 mg of the S-Sepharose-purified unfolded protein described in Example 12b was folded and formulated as follows.

5.2 ml protein solution (equivalent to 20 mg) was diluted with 4.8 ml cold 20 mM Tris/8M urea/50 mM 2-mercaptoethamol/1M NaCl pH8.5, to yield 10 ml of a 2.0 mg/ml solution.

The 10 ml was diluted rapidly with constant stirring into 310 ml freshly prepared, cold (approx. 0-4°C) 20 mM ethanolamine. The solution was left static at 0-4°C for 1 h. Then reduced glutathione was added to 1 mM (by addition of 2.56 ml 125 mM GSH). Then oxidised glutathione was added to 0.5 mM (by addition of 3.2 ml 50 mM GSSG). The solution was left static, in the cold room (~ 2-3°C), for a further 48 h. The solution was then ultrafiltered using a stirred cell and a YM10 (n minal 10,000 molecular weight out-off) membrane to approximately 2 ml. The

solution was clear. The ultrafiltration cell was washed with approximately 2 ml 20 mM ethanolamine and the wash and the ultrafiltered retentate were pooled to give a final volume of 3.7 ml.

2.2 ml of this solution was buffer-exchanged into 3.2 ml 50 mM formic acid using Sephadex G25 (PD10). The buffer-exchanged material was regarded as the product, and it was stored at -40°C. Analysis of an aliquot of the product showed it contained 1.6 mg protein/ml, that by SDS PAGE under non-reducing conditions a single major band of M_r~28,000 was present and that N-terminal sequencing of the band (MQXNAPE) was consistent with the expected sequence. In addition the preparation gave an IH_{50} value (see Method (xvi)) of approximately 1 µg/ml.

IN THE FIGURES

Fig. 1 Plasmid pBROC413. *bla* indicates the ampicillin resistance gene, ø10 the T7 RNA polymerase promoter and rbs the ribosome binding site. Arrows for ø10 and *bla* give the direction of transcription. The polylinker site has been indicated. The plasmid is not drawn to scale and the size is approximate.

Figure 2 illustrates the construction from pDB1010-D11 and pBROC435 of plasmid pDB1013-5-4 coding for SCR 1 + 2 + 3. Plasmid sizes are approximate and are not drawn to scale.

Figure 3 illustrates the construction from pDB1013-5-4 of pDB1018 coding for SCR 1+2+3+4. Plasmid sizes are approximate and are not drawn to scale.

REFERENCES USED IN EXAMPLES OR GENERAL METHODS

1. Chen G-F. and Inouye M. (1990). Suppression of the negative effect of minor arginine codons on gene expression; preferential usage of minor codons within the first 25 codons of the *E.coli* genes. *Nuc.Acids.Res. 18 (6)*: 1465-1473.

2. Dacie, J.V. & Lewis S.M., (1975) *Practical haematology* Fifth Edition, Ed. Churchill Livingstone, Edinburgh & New York pp 3 - 4

3. Devereux, Haeberli and Smithies (1984). A comprehensive set of sequence analysis programmes for the vax. *Nuc.Acids.Res. 12(1)*: 387 - 395

4. Hourcade D., Miesner D. R., Atkinson J. P. & Holers V. M. (1988). Identification of an alternative polyadenylation site in the human C3b/C4b receptor (complement receptor type 1) transcriptional unit and prediction of a secreted form of complement receptor type 1. *J. Exp. Med.* **168** 1255 - 1270

5. Low B. (1968). *Proc.Natl.Acad.Sci.USA 60*: 160

6. Sambrook J., Fritsch E.F. and Maniatis J. (1989). Molecular Cloning: A Laboratory Manual 2nd Edition. Cold Spring Harbour Laboratory Press

7. Studier F.W. and Moffat B.A. (1986). *J.Mol.Biol 189*: 113

8. Tabor S. (1990). Expression using the T7 RNA polymerase/promoter system. In Current Protocols in Molecular Biology (F.A. Ausubel, R. Brent, R.E. Kingston, .D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl, eds) pp. 16.2.1 - 16.2.11 Greene Publishing and Wiley-Interscience, New York

TABLE 1

OLIGO 1 = SEQ ID NO: 1
OLIGO 2 = SEQ ID NO: 2
OLIGO 3 = SEQ ID NO: 3
OLIGO 4 = SEQ ID NO: 4
OLIGO 5 = SEQ ID NO: 5
OLIGO 6 = SEQ ID NO: 6
OLIGO 7 = SEQ ID NO: 7
OLIGO 8 = SEQ ID NO: 8
OLIGO 9 = SEQ ID NO: 9
OLIGO 10 = SEQ ID NO: 10
OLIGO 11 = SEQ ID NO: 11
OLIGO 12 = SEQ ID NO: 12
OLIGO 13 = SEQ ID NO: 13
OLIGO 14 = SEQ ID NO: 14
OLIGO 15 = SEQ ID NO: 15
OLIGO 16 = SEQ ID NO: 16
OLIGO 17 = SEQ ID NO: 17
OLIGO 18 = SEQ ID NO: 18
OLIGO 19 = SEQ ID NO: 19
OLIGO 20 = SEQ ID NO: 20
OLIGO 21 = SEQ ID NO: 21
OLIGO 22 = SEQ ID NO: 22
OLIGO 23 = SEQ ID NO: 23
OLIGO 24 = SEQ ID NO: 24
OLIGO 25 = SEQ ID NO: 25
OLIGO 26 = SEQ ID NO: 26

TABLE 2 Amin acid sequences of proteins, deduced from the cDNA constructs.

The full deduced sequence of the proteins of the Examples are given as follows:

MQ1->K196 of CR-1 is given in SEQ ID NO: 27

MR122->K196 of CR-1 is given in SEQ ID NO: 28

MQ1-S253 of CR-1 is given in SEQ ID NO: 29

The R235H mutant of MQ1-S253 of CR-1 is given in SEQ ID NO: 30

MR122-S253 of CR-1 is given in SEQ ID NO: 31.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: SmithKline Beecham p.l.c. -, - -

(ii) TITLE OF INVENTION: Novel Compounds

(iii) NUMBER OF SEQUENCES: 31

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: SmithKline Beecham Corporate Patents
(B) STREET: Great Burgh, Yew Tree Bottom Road
(C) CITY: Epsom
(D) STATE: Surrey
(E) COUNTRY: England
(F) ZIP: KT18 5XQ

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Valentine, Jill B
(B) REGISTRATION NUMBER: G.A. 26758
(C) REFERENCE/DOCKET NUMBER: P30423

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 0737364158

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 87 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TATGCAGTGC AACGCTCCGG AATGGCTGCC GTTCGCGCGC CCGACCAACC TGACTGATGA 60

ATTTGAGTTC CCGATCGGTA CCTACCT 87

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 93 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CGTAGTTCAG GTAGGTACCG ATCGGGAACT CAAATTCATC AGTCAGGTTG GTCGGGCGCG 60

CGAACGGCAG CCATTCCGGA GCGTTGCACT GCA 93

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 101 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAACTACGAA TGCCGCCCGG GTTATAGCGG CCGCCCGTTT TCTATCATCT GCCTGAAAAA 60

CTCTGTCTGG ACTGGTGCTA AGGACCGTTG CCGACGTAAA T 101

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 101 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ACGACAAGAT TTACGTCGGC AACGGTCCTT AGCACCAGTC CAGACAGAGT TTTTCAGGCA 60

GATGATAGAA AACGGGCGGC CGCTATAACC CGGGCGGCAT T 101

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 101 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTTGTCGTAA TCCGCCAGAT CCGGTTAACG GCATGGTGCA TGTGATCAAA GGCATCCAGT 60

TCGGTTCCCA AATTAAATAT TCTTGTACTA AAGGTTACCG T 101

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 101 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCAATCAGAC GGTAACCTTT AGTACAAGAA TATTTAATTT GGGAACCGAA CTGGATGCCT 60

TTGATCACAT GCACCATGCC GTTAACCGGA TCTGGCGGAT T 101

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 94 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTGATTGGTT CCTCCAGCGC TACATGCATC ATCTCTGGTG ATACTGTCAT TTGGGATAAT 60

GAAACACCGA TTTGTGACCG AATTCAGTAA TAAA 94

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 90 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGCTTTTATT ACTGAATTCG GTCACAAATC GGTGTTTCAT TATCCCAAAT GACAGTATCA 60

CCAGAGATGA TGCATGTAGC GCTGGAGGAA 90

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 72 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TATGCGAATT CCGTGTGGTC TGCCGCCGAC CATCACCAAC GGTGATTTCA TCTCTACCAA 60

TCGCGAGAAT TT 72

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 78 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CATAGTGAAA ATTCTCGCGA TTGGTAGAGA TGAAATCACC GTTGGTGATG GTCGGCGGCA 60

GACCACACGG AATTCGCA 78

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 85 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TCACTATGGT TCTGTGGTGA CCTACCGCTG CAATCCGGGT AGCGGTGGTC GTAAGGTGTT 60

TGAGCTCGTG GGTGAGCCGT CCATC 85

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 85 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GTGCAGTAGA TGGACGGCTC ACCCACGAGC TCAAACACCT TACGACCACC GCTACCCGGA 60

TTGCAGCGGT AGGTCACCAC AGAAC 85

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 79 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TACTGCACTA GTAATGACGA TCAAGTGGGC ATCTGGAGCG GCCCGGCACC GCAGTGCATC 60

ATCCCGAACA AATAATAAA 79

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 75 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AGCTTTTATT ATTTGTTCGG GATGATGCAC TGCGGTGCCG GGCCGCTCCA GATGCCCACT 60

TGATCGTCAT TACTA 75

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GAGACTCATA TGCAGTGCAA CGCTCCGGAA 30

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GTCAGCAAGC TTTTATTACT GAATTCGGTC 30

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ATCGTAACTA GTAACGACGA TCAAGTGGGC 30

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ATGACTAAGC TTTTATTATG AGCAGCTCGG 30

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TAACTAGTAA CGACGATCAA GTGGGCATCT GG 32

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CTAAGCTTTT ATTATGAGCA GCTCGGGAGT TCC 33

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 81 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CTAGTAACGA CGATCAAGTG GGCATCTGGA GCGGCCCGGC ACCGCAGTGC ATCATCCCGA 60

ACAAATGCAC GCCGCCAAAT G 81

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 85 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GTTCTCCACA TTTGGCGGCG TGCATTTGTT CGGGATGATG CACTGCGGTG CCGGGCCGCT 60

CCAGATGCCC ACTTGATCGT CGTTA 85

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 90 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TGGAGAACGG TATCCTGGTA TCTGACAACC GTTCTCTGTT CTCTTTAAAC GAAGTTGTAG 60

AGTTTCGTTG TCAGCCGGGC TTTGTTATGA 90

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 90 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CGGACCTTTC ATAACAAAGC CCGGCTGACA ACGAAACTCT ACAACTTCGT TTAAAGAGAA 60

CAGAGAACGG TTGTCAGATA CCAGGATACC 90

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 72 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

AAGGTCCGCG CCGTGTGAAG TGCCAGGCCT TGAACAAATG GGAGCCGGAA CTCCCGAGCT 60

GCTCATAATA AA 72

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AGCTTTTATT ATGAGCAGCT CGGGAGTTCC GGCTCCCATT TGTTCAAGGC CTGGCACTTC 60

ACACGGCG 68

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 197 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Met Gln Cys Asn Ala Pro Glu Trp Leu Pro Phe Ala Arg Pro Thr Asn
1 5 10 15

Leu Thr Asp Glu Phe Glu Phe Pro Ile Gly Thr Tyr Leu Asn Tyr Glu
20 25 30

```
Cys Arg Pro Gly Tyr Ser Gly Arg Pro Phe Ser Ile Ile Cys Leu Lys
        35                  40                  45

Asn Ser Val Trp Thr Gly Ala Lys Asp Arg Cys Arg Arg Lys Ser Cys
    50                  55                  60

Arg Asn Pro Pro Asp Pro Val Asn Gly Met Val His Val Ile Lys Gly
65                  70                  75                  80

Ile Gln Phe Gly Ser Gln Ile Lys Tyr Ser Cys Thr Lys Gly Tyr Arg
                85                  90                  95

Leu Ile Gly Ser Ser Ser Ala Thr Cys Ile Ile Ser Gly Asp Thr Val
            100                 105                 110

Ile Trp Asp Asn Glu Thr Pro Ile Cys Asp Arg Ile Pro Cys Gly Leu
        115                 120                 125

Pro Pro Thr Ile Thr Asn Gly Asp Phe Ile Ser Thr Asn Arg Glu Asn
    130                 135                 140

Phe His Tyr Gly Ser Val Val Thr Tyr Arg Cys Asn Pro Gly Ser Gly
145                 150                 155                 160

Gly Arg Lys Val Phe Glu Leu Val Gly Glu Pro Ser Ile Tyr Cys Thr
                165                 170                 175

Ser Asn Asp Asp Gln Val Gly Ile Trp Ser Gly Pro Ala Pro Gln Cys
            180                 185                 190

Ile Ile Pro Asn Lys
        195
```

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 76 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

```
Met Arg Ile Pro Cys Gly Leu Pro Pro Thr Ile Thr Asn Gly Asp Phe
1               5                   10                  15

Ile Ser Thr Asn Arg Glu Asn Phe His Tyr Gly Ser Val Val Thr Tyr
            20                  25                  30

Arg Cys Asn Pro Gly Ser Gly Gly Arg Lys Val Phe Glu Leu Val Gly
        35                  40                  45

Glu Pro Ser Ile Tyr Cys Thr Ser Asn Asp Asp Gln Val Gly Ile Trp
    50                  55                  60

Ser Gly Pro Ala Pro Gln Cys Ile Ile Pro Asn Lys
65                  70                  75
```

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 254 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```
Met Gln Cys Asn Ala Pro Glu Trp Leu Pro Phe Ala Arg Pro Thr Asn
1               5                   10                  15

Leu Thr Asp Glu Phe Glu Phe Pro Ile Gly Thr Tyr Leu Asn Tyr Glu
            20                  25                  30

Cys Arg Pro Gly Tyr Ser Gly Arg Pro Phe Ser Ile Ile Cys Leu Lys
        35                  40                  45

Asn Ser Val Trp Thr Gly Ala Lys Asp Arg Cys Arg Arg Lys Ser Cys
```

```
    50                  55                  60

Arg Asn Pro Pro Asp Pro Val Asn Gly Met Val His Val Ile Lys Gly
65                  70                  75                  80

Ile Gln Phe Gly Ser Gln Ile Lys Tyr Ser Cys Thr Lys Gly Tyr Arg
                85                  90                  95

Leu Ile Gly Ser Ser Ser Ala Thr Cys Ile Ile Ser Gly Asp Thr Val
            100                 105                 110

Ile Trp Asp Asn Glu Thr Pro Ile Cys Asp Arg Ile Pro Cys Gly Leu
        115                 120                 125

Pro Pro Thr Ile Thr Asn Gly Asp Phe Ile Ser Thr Asn Arg Glu Asn
    130                 135                 140

Phe His Tyr Gly Ser Val Val Thr Tyr Arg Cys Asn Pro Gly Ser Gly
145                 150                 155                 160

Gly Arg Lys Val Phe Glu Leu Val Gly Glu Pro Ser Ile Tyr Cys Thr
                165                 170                 175

Ser Asn Asp Asp Gln Val Gly Ile Trp Ser Gly Pro Ala Pro Gln Cys
            180                 185                 190

Ile Ile Pro Asn Lys Cys Thr Pro Pro Asn Val Glu Asn Gly Ile Leu
        195                 200                 205

Val Ser Asp Asn Arg Ser Leu Phe Ser Leu Asn Glu Val Val Glu Phe
    210                 215                 220

Arg Cys Gln Pro Gly Phe Val Met Lys Gly Pro Arg Arg Val Lys Cys
225                 230                 235                 240

Gln Ala Leu Asn Lys Trp Glu Pro Glu Leu Pro Ser Cys Ser
                245                 250
```

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 254 amino acids

(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

```
Met Gln Cys Asn Ala Pro Glu Trp Leu Pro Phe Ala Arg Pro Thr Asn
1               5                   10                  15

Leu Thr Asp Glu Phe Glu Phe Pro Ile Gly Thr Tyr Leu Asn Tyr Glu
            20                  25                  30

Cys Arg Pro Gly Tyr Ser Gly Arg Pro Phe Ser Ile Ile Cys Leu Lys
        35                  40                  45

Asn Ser Val Trp Thr Gly Ala Lys Asp Arg Cys Arg Arg Lys Ser Cys
    50                  55                  60

Arg Asn Pro Pro Asp Pro Val Asn Gly Met Val His Val Ile Lys Gly
65                  70                  75                  80

Ile Gln Phe Gly Ser Gln Ile Lys Tyr Ser Cys Thr Lys Gly Tyr Arg
                85                  90                  95

Leu Ile Gly Ser Ser Ser Ala Thr Cys Ile Ile Ser Gly Asp Thr Val
            100                 105                 110

Ile Trp Asp Asn Glu Thr Pro Ile Cys Asp Arg Ile Pro Cys Gly Leu
        115                 120                 125

Pro Pro Thr Ile Thr Asn Gly Asp Phe Ile Ser Thr Asn Arg Glu Asn
    130                 135                 140

Phe His Tyr Gly Ser Val Val Thr Tyr Arg Cys Asn Pro Gly Ser Gly
145                 150                 155                 160

Gly Arg Lys Val Phe Glu Leu Val Gly Glu Pro Ser Ile Tyr Cys Thr
                165                 170                 175
```

```
Ser Asn Asp Asp Gln Val Gly Ile Trp Ser Gly Pro Ala Pro Gln Cys
            180                 185                 190

Ile Ile Pro Asn Lys Cys Thr Pro Pro Asn Val Glu Asn Gly Ile Leu
        195                 200                 205

Val Ser Asp Asn Arg Ser Leu Phe Ser Leu Asn Glu Val Val Glu Phe
    210                 215                 220

Arg Cys Gln Pro Gly Phe Val Met Lys Gly Pro His Arg Val Lys Cys
225                 230                 235                 240

Gln Ala Leu Asn Lys Trp Glu Pro Glu Leu Pro Ser Cys Ser
                245                 250
```

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 133 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```
Met Arg Ile Pro Cys Gly Leu Pro Pro Thr Ile Thr Asn Gly Asp Phe
1               5                   10                  15

Ile Ser Thr Asn Arg Glu Asn Phe His Tyr Gly Ser Val Val Thr Tyr
            20                  25                  30

Arg Cys Asn Pro Gly Ser Gly Gly Arg Lys Val Phe Glu Leu Val Gly
        35                  40                  45

Glu Pro Ser Ile Tyr Cys Thr Ser Asn Asp Asp Gln Val Gly Ile Trp
    50                  55                  60

Ser Gly Pro Ala Pro Gln Cys Ile Ile Pro Asn Lys Cys Thr Pro Pro
65                  70                  75                  80
```

```
Asn Val Glu Asn Gly Ile Leu Val Ser Asp Asn Arg Ser Leu Phe Ser
                85                  90                  95

Leu Asn Glu Val Val Glu Phe Arg Cys Gln Pro Gly Phe Val Met Lys
            100                 105                 110

Gly Pro Arg Arg Val Lys Cys Gln Ala Leu Asn Lys Trp Glu Pro Glu
        115                 120                 125

Leu Pro Ser Cys Ser
    130
```

Claims

1. A soluble polypeptide comprising, in sequence, one to four short consensus repeats (SCR) selected from SCR 1, 2, 3 and 4 of long homologous repeat A (LHR-A) as the only structurally and functionally intact SCR domains of CR1 and including at least SCR3.

2. A polypeptide according to claim 1 which comprises, in sequence, SCR 1, 2 and 3 of LHR-A as the only structurally and functionally intact SCR domains of CR1.

3. A polypeptide according to claim 2 of formula (I):

$$NH_2\text{-}V^1\text{-}SCR1\text{-}W^1\text{-}SCR2\text{-}X^1\text{-}SCR3\text{-}Y^1\text{-}OH \qquad (I)$$

in which SCR1 represents residues 2-58 of mature CR1, SCR2 represents residues 63-120 of mature CR1, SCR3 represents residues 125-191 of mature CR1, and V^1, W^1, X^1 and Y^1 represent bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

4. A polypeptide according to claim 3 in which W^1, X^1 and Y^1 represent residues 59-62, 121-124 and 192-196, respectively, of mature CR1 and V^1 represents residue 1 of mature CR1 optionally linked via its N-terminus to methionine.

5. A polypeptide according to claim 1 which comprises, in sequence, SCR 1, 2, 3 and 4 of LHR-A as the only structurally and functionally intact SCR domains of CR1.

6. A polypeptide according to claim 5 of formula (II):

$$NH_2\text{-}V^2\text{-}SCR1\text{-}W^2\text{-}SCR2\text{-}X^2\text{-}SCR3\text{-}Y^2\text{-}SCR4\text{-}Z^2OH \qquad (II)$$

in which SCR1, SCR2 and SCR3 are as defined in claim 3, SCR4 represents residues 197-252 of mature CR1 and V^2, W^2, X^2, Y^2 and Z^2 represents bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

7. A polypeptide according to claim 6 in which W^2, X^2, Y^2 and Z^2 represent residues 59-62, 121-124, 192-196, and residues 253 respectively, of mature CR1 and V^2 represents residue 1 of mature CR1 optionally linked via its N-terminus to methionine.

8. A polypeptide according to claim 7 in which arginine 235 is replaced by histidine.

9. A polypeptide according to claim 1 of formula (III):

$$NH_2\text{-}X^3\text{-}SCR3\text{-}Y^3\text{-}OH \qquad (III)$$

in which SCR3 is as defined in claim 3 and X^3 and Y^3 represent bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

10. A polypeptide according to claim 9 in which X^3 represents amino acids 122-124 of mature CR1 optionally linked to methionine at its N-terminus and Y^4 represents amino acids 192-196 of mature CR1.

11. A polypeptide according to claim 1 of formula (IV):

$$NH_2\text{-}X^4\text{-}SCR3\text{-}Y^4\text{-}SCR4\text{-}Z^4\text{-}OH \qquad (IV)$$

in which SCR3 and SCR4 are as defined in claim 6 and X^4, Y^4 and Z^4 represent bonds or short linking sequences of amino acids, preferably 1 to 5 residues in length and which are preferably derived from native interdomain sequences in CR1.

12. A polypeptide according to claim 11 in which X^4 represents amino acids 122-124 of mature CR1 optionally linked to methionine at its N-terminus and Y^4 and Z^4 represent amino acids 192-196 and 253 respectively of mature CR1.

13. A polypeptide having the amino acid sequence given in SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30 or SEQ ID NO:31.

14. A DNA polymer comprising a nucleotide sequence that encodes the polypeptide according to claim 1.

15. A replicable expression vector capable, in a host cell, of expressing the DNA polymer of claim14.

16. A host cell transformed with the replicable expression vector of claim 15.

17. A process for preparing a CR1 polypeptide according to claim 1 which process comprises expressing DNA encoding said polypeptide in a recombinant host cell and recovering the product.

18. A pharmaceutical composition comprising a therapeutically effective amount of a polypeptide according to claim 1 and a pharmaceutically acceptable carrier or excipient.

19. A polypeptide according to claim 1 for use as an active therapeutic substance.

20. A method of treating a disease or disorder associated with inflammation or inappropriate complement activation comprising administering to a subject in need of such treatment a therapeutically effective amount of a polypeptide according to claim 1.

21. The use of a polypeptide of claim 1 in the manufacture of a medicament for the treatment of a disease or disorder associated with inflammation or inappropriate complement activation.

*****

The following methods and examples further illustrate aspects of the invention.

Introduction of DNA into *E. coli*

Plasmids were transformed into *E. coli* XL1-Blue (Stratagene), HMS174(DE3) (Novagen, UK) or UT5600(DE3) (see below) that had been made competent using calcium chloride as described in Sambrook *et al*, (op.cit.). UT5600 was purchased from New England Biolabs (#801-I) and was converted to a DE3 lysogen. UT5600 was isolated as a mutant of K12 strain RW193 (itself derived from AB1515) which was insensitive to colicin -B (McIntosh et al. (1979) J.Bact. 137 p653). It was not initially known that *ompT* had been lost, but further work by the same group showed that protein *a* (now *OmpT*) was lacking (Earhart et al (1979) FEMS Micro Letts 6 p277). The nature of the mutation was determined to be a large deletion (Elish et al (1988) J Gen Micro **134** 1355).

DNA sequencing

DNA sequencing was contracted out to Lark (Saffron Walden, Essex UK) or MWG (Milton Keynes, UK).

Production of oligonucleotides

Oligonucleotides were purchased from Cruachem (UK) or Genosys-Sigma (Pampisford, Cambridgeshire UK)

Polymerase chain reaction amplification of DNA

Purified DNA or DNA fragments from ligation reactions or DNA fragments excised and purified from agarose gels were amplified by PCR from two primers complementary to the 5' ends of the DNA fragment. Approximately 0.1 – 1 mg of DNA was mixed with commercially available buffers for PCR amplification such as 10 mM Tris pH 8.3 (at 25^{o}C), 50 mM KCl, 0.1% gelatin; $MgCl_2$ concentrations were varied from 1.5 mM to 6 mM to find a suitable concentration for each reaction. Oligonucleotide primers were added to a final concentration of 2 mM; each dNTP was added to a final concentration of 0.2 mM. 1 unit of Taq DNA polymerase was then added to the reaction mixture (purchased from a commercial source, e.g. Gibco). The final reaction volume varied from 20 ml to 100 ml, which was overlayed with mineral oil to prevent evaporation. Thermal

CIP
new

cycling was then started on a thermal cycler such as the PCR machine from MJ Research. A typical example of conditions used was 94°C for 5 minute, 55°C for 1 minute, and 72 °C for 2 minutes; however, the optimal temperatures for cycling can be determined empirically by workers skilled in the art. The DNA fragment was amplified by repeating this temperature cycle for a number of times, typically 30 times.

pET15b vector for DAF expression

The pET15b expression vector is a T7 promotor based vector available commercially through Novagen (Wisconsin, USA). Briefly, the vector carries the following features which make it a useful vehicle for the expression of heterologous proteins in E. coli; a selectable antibiotic marker (β-lactamase) conferring ampicillin resistance, a copy of the lacI gene providing lac repression in strains of E. coli that are lacI^{-}, and the T7-lac promoter. The T7-lac promoter combines the T7 RNA polymerase promoter sequences with the lacI repressor binding site from the E. coli lactose operon. This reduces expression of the cloned gene in the absence of the inducer isopropyl β-D thiogalactopyranoside (IPTG). Downstream of the T7 promoter is a multiple cloning site built into a region of sequence which codes for a polyhistidine tag sequence. Translation initiates at the methionine codon at position 332-330 of the vector sequence and proceeds counter-clockwise to yield the following peptide:
SEQ ID 65
MGSSHHHHHHSSGLVPRGSH. The six histidine residues allow for purification of the fusion protein by metal chelation chromatography, whilst the GLPVR motif constitutes a thrombin cleavage site for removal of the peptide from the fusion protein after purification. Three restriction enzyme sites are provided for the insertion of cloned DNA in-frame with the polyhistidine leader. These are NdeI (CATATG), XhoI (CTCGAG) and BamHI (GGATCC). Use of the NdeI site to overlap the methionine initiation codon of the cloned gene removes the possibility of unwanted amino acids at the N-terminus of the fusion protein. At the 3' end of the multiple cloning site is the T7 transcriptional terminator.

Colorometric determination of protein concentration

Protein concentration determination utilised a colorometric method utilising Coomassie Plus Protein Assay Reagent (Pierce Chemical Company) according to

the manufacturer's instructions. The assay used a reference standard of the protein of Example 6.

Identification of proteins by Western blot

For certain procedures, it is necessary to characterise the expression of recombinant proteins by an immunological method termed a Western blot. In this method, proteins to be analysed are separated by SDS-PAGE, transferred to a protein binding membrane such as polyvinylidene difluoride (PVDF), and then probed with an antibody that is specific for the target protein. Typically, the binding of the first antibody is detected by the addition of an enzyme-labelled secondary antibody and an appropriate solution which contains a chromogenic substrate. One procedure for the transfer of proteins to a protein-binding membrane was as follows. After SDS-PAGE, the proteins on the gel were transferred by electrotransfer to a protein-binding surface such PVDF. In this procedure, two sheets of filter paper (3M, Whatman) soaked in 0.3M Tris, 10% (v/v) methanol, pH10.4, were placed on the anode of an electroblotter (Semi-dry blotter, Biorad). These filter papers were then overlayed by a further two sheets of filter paper soaked in 25mM Tris, 10% (v/v) methanol, pH10.4. On top of this stack of filter papers was placed a sheet of PVDF membrane which had been pre-wetted in methanol and then soaked in a buffer that comprises 25mM Tris, 10% (v/v) methanol, pH10.4. The SDS-PAGE gel was then placed on the top of the PVDF membrane, and overlayed with two sheets of filter paper soaked in 25mM Tris, 192 mM 6-amino-n-caproic acid, 10% (v/v) methanol. The cathode of the electroblotter was then placed on top of the stack of filter papers, gel and membrane, and the proteins transferred by passing a current between the electrodes at 15V for 30 minutes. Subsequent steps for the detection of the transferred proteins were described in the Novex WesternBreeze System (Invitrogen). For the detection of human CD59, a rat anti-CD59 monoclonal antibody YTH53.1 (Davies *et al., J. Exp. Med.* **170**, 637, 1989) was used together with an enzyme-labeled anti-rat secondary antibody. For the identification of His-tagged DAF, an anti his-tag monoclonal antibody was used.

Purification of CD59 from human urine

Urine was collected into 10mM sodium azide/5mM benzamidine over approximately 48hrs. The urine was then passed through a fluted coarse filter to

remove aggregates. The urine was then concentrated to approximately 150mls in a Pellicon concentrator fitted with a membrane cassette with a 10kDa MW cut-off membrane. Insoluble material was removed by centrifugation at 10000xg for 30 minutes. The supernatant was then applied to a CNBr-activated Sepharose 4B affinity column prepared with the rat monoclonal anti-CD59 antibody YTH 53.1 (Davies *et al. J. Exp. Med.* **170**, 637, 1989). The column was washed overnight with 1M NaCl and bound material eluted with 4M $MgCl_2$. The protein content of each 1ml fraction eluted from the column was determined by measuring absorbance at OD280nm. The fractions containing the most protein were then pooled and dialysed through a 10kDa MW cut off membrane into a solution containing 0.9% NaCl, and then dialysed by a similar procedure into PBS. The dialysed protein was then concentrated using a stirred cell ultrafiltration device (Amicon) fitted with a 10kDa MW cut-off membrane. The material may be further purified by gel filtration in 10mM Hepes, 140mM NaCl, pH7.4, on a Superdex S-75 fast protein liquid chromatography system (Pharmacia) or Sephadex G-75. This method gave a yield of around 7mg pure protein from 20L urine.

Expression and purification of recombinant soluble CD59 from CHO cells

Soluble CD59 was expressed in a recombinant form from Chinese Hamster Ovary cells as follows. Briefly, the polymerase chain reaction was used to produce a truncated cDNA encoding soluble CD59 from a full length cDNA (Davies *et al. J. Exp. Med.* **170**, 637, 1989). A mutation was introduced into the cDNA at codon 18 of the mature protein which changed the Asn codon for Ala. The procedure for this site-directed mutagenesis can be performed by a number of methods including the Quickchange mutagenesis kit (Stratagene). To introduce the modified gene into the CHO expression plasmid pDR2EF1alpha, the polymerase chain reaction was used with two oligonucleotides; the first oligonucleotide was complementary to the first seven codons at the N-terminus of the mature CD59 protein; and the 3' oligonucleotide introduces a termination codon immediately following the codon for Asn-70 of the CD59 cDNA. These oligonucleotides were also designed to contain recognition sequences for restriction endonucleases compatible with the polylinker site of the CHO expression vector. The DNA fragment resulting from the PCR amplification was ligated into a CHO expression

vector and this plasmid transfected with calcium phosphate into CHO cells. Cells that had become stabily transfected were selected from untransfected cells by growth in medium that contained the antibiotic hygromycin. Individual transformants were picked and for each clone the expression of CD59 was analysed by ELISA. The highest expressing clone was chosen for large-scale production of CD59 using a variety of techniques including the use of cell factories (Nunc).

To purify the CD59, the culture medium was precleared by centrifugation at 10000xg for 30 minutes. The soluble CD59 was then purified using an immunoaffinity column containing the monoclonal antibody YTH53.1 (Davies *et al. J. Exp. Med.* **170**, 637, 1989), as described above. The protein was then stored in PBS at concentrations of up to 5mg/mL at –70 ^{o}C .

Preparation of C56 euglobulin

C56 euglobulin was an essential reagent that was used for the C5b6-initiated reactive lysis of erythrocytes. C56 euglobulin can be generated in and purifed from some acute-phase sera from post-trauma individuals (such as sports injuries, surgery or childbirth). Blood was drawn from donors in the acute phase of inflammation and allowed to clot at room temperature. To each 10mls of serum, 0.5mls of yeast suspension was added and the mixture incubated overnight on a rotator at room temperature. The serum was centrifuged to remove the yeast and dialysed against 0.02M Na/K phosphate, pH 5.4. The precipitate (containing the C56 euglobulin) was collected by centrifugation and redisolved in 0.01M Na/K phosphate/0.05M NaCl, pH7.0 containing 25% v/v glycerol.

C5b6-initiated reactive lysis of erythrocytes

Guinea pig erythrocytes (TCS Microbiological, UK) were washed twice in PBS and resuspended to 5% by volume in PBS/0.05% CHAPS. 50 ml of these cells were placed in the wells of a round-bottomed microtitre plate. Samples to be tested were diluted in PBS/0.05% CHAPS and 50 ml of these test solutions added to the wells containing the guinea pig erythrocytes. The plate was then incubated at 37 ^{o}C for 20 minutes to allow binding of the samples to the erythrocytes. The microtitre plates were then centrifuged at 1000rpm for 5 minutes to pellet the cells using a benchtop centrifuge. The supernatants were removed and the cell pellets

resuspended in 50 ml PBS/10mM EDTA. To this cell suspension was added 10 ml of a C56 euglobulin solution that varied in concentration in different experiments from between 1:50 to 1:500 dilution in PBS/10mM EDTA. This solution was mixed with the cells by placing the microtitre plate on a microtitre plate shaker for 2 minutes. To this solution was then added a 90 ml of a dilution of normal human serum (from 1:50 to 1:500 in PBS/10mM EDTA). The solutions were mixed by placing the microtitre plate on a plate shaker for a further 2 minutes. The plate was then incubated at 37 °C for 30 minutes. To determine the degree of haemolysis, the plate was then placed in a benchtop centrifuge and spun at 1800 rpm for 3 minutes. 100 ml of the supernatant was transferred to a clear flat bottomed microtitre plate and the absorbance at 410nm measured spectroscopically. As controls, guinea pig erythrocytes were treated in an identical manner to the test samples with the following exceptions. In the first stage of the assay, the control samples were incubated with 50 ml of PBS/10mM EDTA for 20 minutes at 37 °C. After centrifugation, a spontaneous lysis control was prepared by resuspending the cells in 150 ml PBS/10mM EDTA; by contrast, for the maximum lysis control, the cells were resuspended in 150 ml water.

Brief Overview of Examples 25 to 36

Example 25: Synthesis and characterization of a membrane- targeted derivative of soluble human urinary CD59 (APT632).

Example 26: Synthesis and characterization of a membrane- targeted derivative of human recombinant soluble CD59 (APT637).

Example 27: An alternative Method for the production of urinary (APT2047) and recombinant (APT2059) human CD59 membrane- targeted derivatives using linkage through protein carbohydrate.

Example 28: A method for the preparation of recombinant human CD59 with a C-terminal cysteine, expressed in yeast (APT633).

Example 29: A method for the preparation of recombinant human CD59 with a C-terminal cysteine, expressed in *E. coli* (APT635).

Example 30: A method for the preparation of recombinant human CD59 with a C-terminal cysteine, expressed in baculovirus/insect cells (APT2060).

Example 31: A method for the preparation of recombinant human CD59 with a C-terminal cysteine, expressed in Chinese hamster ovary cells (APT2061).

Example 32: A Method for the conjugation of the membrane-localising agent APT542 to APT633, APT635, APT2060 or APT2061.

Example 33: Synthesis and characterization of APT2057 (Human DAF short consensus repeats 2-4).

Example 34: Synthesis and characterization of APT2058 (Human DAF short consensus repeats 1-4).

Example 35: Synthesis and characterization of APT2160 (APT2058 derivatized with APT542).

Example 36: Synthesis and characterization of APT2184 (APT2057 derivatized with APT542).

Example 25: Synthesis and characterization of urine-derived CD59 linked to MSWP-1 (APT632)

APT632 was synthesized in two steps from soluble CD59 isolated from human urine (APT631; SEQ. ID NO: 37) as described in Methods. APT631 in PBS (200 μL of a 1.9mg/mL solution) was mixed with 2-iminothiolane (2 μL of a 100mM solution) and the mixture incubated at room temperature for 30 minutes. The solution was then dialysed into PBS to remove unreacted 2-iminothiolane, and a solution of tris-2-carboxyethyl phosphine (4μL of a 10mM solution in 10mM Hepes, pH7.4) added, and the mixture left overnight at room temperature. To this solution, 10μL of APT542 (MSWP-1; 21mM in dimethyl sulphoxide; SEQ. ID NO. 38) was added and incubated at room temperature for 2 h. The product APT632 was characterized by the appearance of a protein species that migrated at approximately 21kDa as analysed by SDS-PAGE. A reactive lysis assay (described in Methods) demonstrated that APT632, at concentrations greater than 0.5nM, protected guinea pig erythrocytes from complement-mediated lysis by a 1:100 dilution of human serum; by contrast, no significant protection from lysis was observed with the unmodified form (APT631).

Example 26: Synthesis and characterization of recombinant CD59 produced in CHO cells linked to MSWP-1 (APT637)

APT637 was synthesized in two steps from soluble human CD59 that is expressed in a recombinant form from chinese hamster ovary cells (APT634; SEQ ID NO: 39). APT634 in PBS (200 μL of a 300μM solution) was mixed with 2-iminothiolane (6 μL of a 10mM solution) and the mixture incubated at room temperature for 30 minutes. The solution was then dialysed into PBS to remove unreacted 2-iminothiolane, and a solution of tris-2-carboxyethyl phosphine (4μL of a 10mM solution in 10mM Hepes, pH7.4) added, and the mixture left overnight at room temperature. To this solution, 10μL of APT542 (21mM in dimethyl sulphoxide) was added and incubated at room temperature for 2 h. The product APT637 was characterized by the appearance of a protein species which migrated at approximately 10kDa as analysed by SDS-PAGE as described in methods. A reactive lysis assay (described in Methods) demonstrated that APT637, at concentrations greater than 0.5nM, protected guinea pig erythrocytes from complement-mediated lysis by a 1:100 dilution of human serum; by contrast, no significant protection from lysis was observed with the unmodified form (APT634).

Example 27: A Method for the Production of CD59 derivatives linked to MSWP-1 *via* a carbohydrate linkage (APT2047 and APT2059)

APT2047 is a conjugate of APT634 (SEQ ID NO: 39) and APT542 (SEQ ID NO: 38), and APT2059 is a conjugate of APT631 (SEQ ID NO: 37) and APT542, in which the linkage of each pair of compounds is through a modified carbohydrate moeity on the CD59 protein. APT2047 and APT2059 are synthesized in three steps from APT634 or APT631. The first step involves the reaction of the proteins APT634 or APT631 at a concentration of 1mg/ml with 10mM sodium periodate for 1h in the dark, in a solution of 0.1M sodium acetate, pH5.5. To this mixture is added glycerol to a final concentration of 15mM and the solution placed on ice for 5 minutes. The mixture is then dialysed into 0.1M sodium acetate, pH5.5 to remove excess sodium periodate and glycerol. In the second step, the sodium periodate-treated proteins are reacted with a solution of (4-[4-N-

maleimidophenyl]butyric acid hydrazide hydrochloride (MPBH) at a final concentration of 1mg/ml for 2h with stirring. After this procedure, unreacted MPBH is removed by dialysis into a solution of 0.1M phosphate, pH7.0, 50mM NaCl. In the third step of the synthesis, the proteins treated with MPBH are reacted with a solution comprising a 5-fold molar excess of APT544 to CD59 for 2h at room temperature to generate APT2047 and APT2059. The synthesis of these proteins is confirmed by the appearance of a novel proteinaceous species that migrates at approximately 10kDa or 20kDa by SDS-PAGE under non-reducing conditions, respectively. In addition, these proteins protect guinea pig erythrocytes from complement-mediated lysis by human serum at a concentration greater than 0.5nM.

Example 28: A method for the preparation of recombinant human CD59 with a C-terminal cysteine, expressed in yeast (APT633)

APT633 is a protein that comprises soluble human CD59 and a C-terminal cysteine residue following position 81 of the mature CD59 protein. The protein was expressed in a recombinant form in *Pichia pastoris* cells. The polymerase chain reaction was used to produce a truncated cDNA encoding soluble CD59 from a full length cDNA (Davies *et al. J. Exp. Med.* **170**, 637, 1989). The 5' oligonucleotide was complementary to 20 bases of the first 7 codons at the N-terminus of the mature CD59 protein, and the 3' oligonucleotide introduced a cysteine codon and a termination codon immediately following the codon for Ser-81 of the mature CD59 protein. These oligonucleotides were also designed to contain recognition sequences for restriction endonucleases XhoI and EcoRI which are compatible with the polylinker site of the vector pUCPIC (a derivative of pUC19 that contains the alpha-factor leader sequence and multiple cloning site from pPIC9K (Invitrogen). The DNA fragment resulting from the PCR amplification was then ligated into pUCPIC DNA and transformed into the XL1-Blue strain of E. coli (Stratagene). The transfected cells are selected by growth on a petri dish containing LB medium (Sigma) supplemented with ampicillin at a concentration of 100 micrograms/ml (LBAMP). The DNA from single colonies was isolated and sequenced as described in Methods. The DNA that encodes the alpha factor and CD59 was then subcloned into the vector pPIC9K that had been

digested with the restriction endonucleases BamHI and EcoRI. Purified DNA from the resulting plasmid was linearised with the restriction endonuclease PmeI for transformation into P. pastoris strain GS115 (Invitrogen) by spheroplasting according to the manufacturer's instructions. After preliminary selection for clones that are capable of growth on a minimal RD medium(1M sorbitol, 2% w/v dextrose, 1.34% yeast nitrogen base, 4 x 10^{-5} % biotin, 0.005% amino acids) lacking histidine. Clones having undergone multiple integration events were then selected by resistance to the antibiotic geneticin sulphate (G418). Clones that were capable of growth in medium containing G418 at a concentration of 2mg/mL were screened for expression of CD59. Individual colonies were inoculated in 10mL BMG medium (100mM potassium phosphate, pH6.0, 13.4 mg/mL yeast nitrogen base, 0.4 mg/L biotin, 1% (w/v) glycerol) and grown at 30 °C with shaking until clones reached an optical density of 6 as measured spectroscopically at a wavelength of 600nm. The cultures were then transferred to BMM medium (100mM potassium phosphate, pH6.0, 13.4 g/L yeast nitrogen base, 0.4 mg/L biotin, 0.5% methanol) and grown for 48 h at 30 oC with shaking. Culture supernatants were then analysed by SDS-PAGE and Western blot for the presence of APT633 which was observed as a novel proteinaceous species which migrated at approximately 8000Da.

Example 29: A method for the preparation of recombinant human CD59 with a C-terminal cysteine, expressed in *E. coli* (APT635; SEQ ID NO: 41)

APT635 is a protein that comprises soluble human CD59 and a C-terminal cysteine residue following codon 81 of the mature CD59 protein (SEQ ID NO: 41). The protein is expressed in a recombinant form in *E. coli* cells. The polymerase chain reaction was used to produce a truncated cDNA encoding soluble CD59 from a full length cDNA (Davies *et al. J. Exp. Med.* **170**, 637, 1989). The 5' oligonucleotide was complementary to 20 bases of the first 7 codons at the N-terminus of the mature CD59 protein, and the 3' oligonucleotide introduced a cysteine codon and a termination codon immediately following the codon for Ser-81 of the mature CD59 protein. These oligonucleotides were also designed to contain recognition sequences for restriction endonucleases compatible with the polylinker site of pBROC413 (described in WO 94/00571).

The DNA fragment resulting from the PCR amplification was then ligated into pBROC413 DNA and transformed into the UT5600(DE3) strain of *E. coli* (described in Methods). The transfected cells are selected by growth on a petri dish containing LB medium (Sigma) supplemented with ampicillin at a concentration of 100 micrograms/ml (LBAMP). The DNA from single colonies was isolated and sequenced as described in Methods. A single colony representing UT5600(DE3) cells transfected by DNA encoding APT635 was then grown with shaking overnight at 37 oC in LBAMP. This overnight culture was then diluted 1:100 in LBAMP medium and grown with shaking at 37 °C until the culture reached an optical density of 1.0 as determined by absorbance at a wavelength of 600nm. To this culture was added a solution of isopropyl beta-D-thiogalactopyranoside to a final concentration of 1 mM. The culture was then grown for a further 3 hours with shaking at 37°C. The cells are harvested by centrifugation and inclusion bodies isolated as described in WO 94/00571. The expression of APT635 was determined by SDS-PAGE and confirmed by the appearance of a novel protein species that migrated at approximately 8000Da.

Example 30: A method for the preparation of recombinant human CD59 with a C-terminal cysteine, expressed in baculovirus/insect cells (APT2060)

APT2060 is a protein that comprises soluble human CD59 and a C-terminal cysteine residue following codon 81 of the mature CD59 protein (SEQ ID NO: 40) The protein was expressed in a recombinant form in a baculovirus expression system. The polymerase chain reaction was used to produce a truncated cDNA encoding soluble CD59 from a full length cDNA (Davies *et al. J. Exp. Med.* **170,** 637, 1989). The 5' oligonucleotide was complementary to 20 bases of the first 7 codons at the N-terminus of the mature CD59 protein, and the 3' oligonucleotide introduced a cysteine codon and a termination codon immediately following the codon for Ser-81 of the mature CD59 protein. These oligonucleotides were also designed to contain recognition sequences for restriction endonucleases compatible with the polylinker site of pBacPAK 8 baculovirus transfer vector (Clontech). The DNA fragment resulting from the PCR amplification was then ligated into pBacPAK 8 DNA. This plasmid was then transfected into Sf9 cells with Bacfectin (Clontech) and BacPAK6 viral DNA which had been cut with the

restriction endonuclease Bsu36I. This mixture was deposited onto a 50% confluent monolayer of Sf9 cells and left at 28 °C for 3 days. The supernatant was removed and a plaque assay performed on serial dilutions of the transfection supernatant as described in Baculovirus Expression Protocols, Methods in Molecular Biology series, ed. C. Richardson). Individual plaques were then picked into 0.5mL IPL-41 medium (Gibco BRL) containing 1% foetal calf serum. The mixture was left at room temperature for 15 minutes and 100 ml of this solution used to inoculate a 50% confluent monolayer of Sf9 cells. The cells were then left to become infected for 4-5 days at 28°C. After this time, the supernatant was removed and assayed for CD59 expression by Western blot as described in methods. For scale-up of the recombinant virus, the supernatant was used as an inoculum to infect more Sf9 cell monolayers as described above; alternatively, the supernatant can be used to infect Sf9 cells grown in suspension cultures. In this method, 100 mL Sf9 cells at a concentration of 5×10^6 cells/ml in IPL-41 medium containing 1% FCS were inoculated with 50 ml of viral supernatant. The culture was shaken for 5-7 days at 27 °C and cells removed by centrifugation. The recombinant virus may be stored at 4 °C until use. APT2060 may be detected by Western blot as described in Methods and purified using an affinity column as described.

Example 31: A method for the preparation of recombinant human CD59 with a C-terminal cysteine, expressed in Chinese hamster ovary cells (APT2061; SEQ ID. NO: 42)

APT2061 is a protein that comprises soluble human CD59 and a C-terminal cysteine residue at position 71 of the mature protein. The protein may be expressed in a recombinant form in chinese hamster ovary cells as described in Methods. Briefly, the polymerase chain reaction is used to produce a truncated cDNA encoding soluble CD59 from a full length cDNA (Davies *et al. J. Exp. Med.* **170**, 637, 1989). The 5' oligonucleotide is complementary to the first codons at the N-terminus of the mature CD59 protein, and the 3' oligonucleotide introduces a cysteine codon and a termination codon immediately following the codon for Asn-70 of the CD59 cDNA. These oligonucleotides can also designed

to contain recognition sequences for restriction endonucleases compatible with the polylinker site of a CHO expression vector, as described.

Example 32: A Method for the conjugation of APT542 to APT633, APT635, APT2060 or APT2061 to generate compounds APT2062 (see SEQ ID NO: 43), APT2063 (SEQ ID NO: 44), APT2064 (see SEQ ID NO: 43) and APT2065 (SEQ ID NO: 45).

Compounds APT2062, APT2063, APT2064 and APT2065 are generated by treating their parent compounds APT633, APT635, APT2060 and APT2061 with a single molar equivalent of tris-2-carboxyethyl phosphine (TCEP; in 10mM Hepes, pH7.4) overnight at room temperature. To this mixture is added a solution containing 5 molar equivalents of APT542 (MSWP-1) for 2 hours at room temperature.

Example 33: A method for the synthesis and characterization of APT2057 (SEQ ID NO: 46)

APT2057 is a protein that comprises the short consensus repeats 2,3 and 4 of human CD55 (decay accelerating factor, DAF), with a carboxyl terminal cysteine residue and an amino terminal histidine tag motif expressed in a recombinant form in E. coli cells. cDNA to human DAF mRNA was generated from total brain RNA (OriGene Technologies, USA). Reverse transcription was primed with 40 ρmol of primer DAF-R (5'GGAATTCTAAGTCAGCAAGCCCATGGTTACT 3') SEQ ID: 66, 3 μg human brain total RNA and other reagents as recommended by the the RT system manufacturers (Promega, Southampton, UK). Half of the RT reaction (10 μl) was used as template for PCR. Reaction volume was increased to 50 μl by the addition of water, buffer, $MgCl_2$ (to 2 mM), DMSO (to 5%) and 20 ρmol oligonucleotide DAF-F (5'GCATATGACCGTCGCGCGGCCGAGC 3') SEQ ID: 67. One unit of Taq polymerase (MBI Fermentas, Vilnius, Lithuania) was added, and the reaction subjected to 35 cycles of PCR (94°C, 30 sec; 64°C, 30 sec; 72°C, 60 sec). A PCR product of 1156 bp was identified by agarose gel electrophoresis, purified from the gel and ligated using standard procedures into the T-cloning vector pUC57/T (MBI-Fermentas, Vilnius, Lithuania). Positive clones were

identified by PCR screen, analysed by plasmid restriction map and confirmed by full sequence analysis. A plasmid to encode APT2057 was generated by PCR using the pUC-DAF plasmid as template. Primers were designed to amplify the region of the DAF gene encoding amino acids 97-285 (SCR2-4). The 5' primer incorporated an NdeI restriction enzyme site, and a codon specifying glutamine, thereby introducing an amino terminal methionine-glutamine amino acid pair. The 3' primer added a carboxyl terminal cysteine residue and incorporated an EcoRI restriction enzyme site. The PCR product was cloned into the pUC57/T T-vector as described, sequenced, the insert excised with NdeI and EcoRI, and ligated into pET15b (Novagen, Madison, USA, see Methods section). The product of this ligation is the plasmid pET100-02, which expresses DAF(SCR2-4) as an in-frame fusion of a 20 amino acid leader sequence (MGSSHHHHHHSSGLVPRGSH) SEQ ID 65 to the 191 amino acid DAF SCRs2-4. pET100-02 DNA was introduced into E. coli HAMS113 and transformed cells selected by virtue of their ability to grow on LB+agar plates in the presence of 50 μg/ml ampicillin (LBAMP). A single colony representing HAMS113 containing DNA with the coding capacity for APT2057 was grown overnight at 37°C with shaking (200 rpm) in LBAMP medium, then diluted 1:100 into 1 litre fresh LBAMP and growth at 37oC with shaking. Growth was monitored by measurement of culture turbidity at 600nm, and upon reaching an optical density of 0.6, isopropyl β-D thiogalactopyranoside (IPTG)was added to a final concentration of 1 mM, followed by a further 3 hours of growth under the same conditions as described above. The expression of APT2057 was analysed by SDS-PAGE (described in methods). APT2057 appeared as a unique protein product of approximately 24000 Da as estimated by comparative mobility with molecular weight standards. Cells containing APT2057 are harvested by centrifugation and inclusion bodies isolated as follows. Briefly, the cells are resuspended in lysis buffer (50 mM Tris, 1 mM ethylene diamine tetra-acetic acid (ETDA), 50 mM NaCl, pH 8.0) at 50 ml per litre of initial culture. The suspension is lysed by two passages through an Emulsiflex homogensier (Glen-Creston, Middlesex UK), followed by centrifugation at 15000 x g to purify inclusion bodies. Inclusion bodies are initially resuspended to approximately 1 mg.ml^{-1} (as estimated from SDS-PAGE) in 20 mM Tris, 1 mM EDTA, 50mM 2-

mercaptoethanol, pH8.5, and subsequently diluted to a final concentration of 8M urea by the addition of 10 M urea 20 mM Tris, 1 mM EDTA, 50mM 2-mercaptoethanol, pH8.5. This suspension is stirred at 4°C for 16 hours, and insoluble material removed by centrifugation at 15000 x g for 30 minutes. The APT2057 is refolded by 1 in 50 dilution into 20 mM ethanolamine , 1 mM EDTA, pH 11 buffer and static incubation at 4°C for 24 hours. Insoluble material is removed by centrifugation (10000 x g, 10 minutes), and soluble material buffer exchanged into Dulbecco's A PBS, pH 7.4 using an XK50 x 23 cm Sephadex G25 column. Refolded APT2058 is analysed by SDS-PAGE, Western blot and the effectiveness of the protein in a haemolytic assay (described in methods).

Example 34: A method for the synthesis and characterization of APT2058 (SEQ ID NO: 47)

APT2058 is a protein that comprises the short consensus repeats 1,2,3 and 4 of human CD55 (decay accelerating factor, DAF), with a carboxyl terminal cysteine residue and an amino terminal histidine tag motif expressed in a recombinant form in E. coli cells. cDNA to human DAF mRNA was generated from total brain RNA as described in Example 9. A plasmid to encode APT2058 was generated by PCR using the pUC-DAF plasmid as template. Primers were designed to amplify the region of the DAF gene encoding amino acids 35-285 (SCR1-4). The 5' primer incorporated an NdeI restriction enzyme site, and a codon specifying glutamine, thereby introducing an amino terminal methionine-glutamine amino acid pair. The 3' primer added a carboxyl terminal cysteine residue and incorporated an EcoRI restriction enzyme site. The PCR product was cloned into the pUC57/T T-vector as described, sequenced, the insert excised with NdeI and EcoRI, and ligated into pET15b (Novagen, Madison, USA). The product of this ligation is the plasmid pET99-02, which expresses DAF (SCR1-4) as an in-frame fusion of a 20 amino acid leader sequence (MGSSHHHHHHSSGLVPRGSH) SEQ ID: 65 to the 251 amino acid DAF SCRs1-4 (APT2058). pET99-02 DNA was introduced into E. coli HAMS113 (see methods) and expression of the recombinant protein induced as described in Example 1. The expression of APT2058 was analysed by SDS-PAGE (described in methods). APT2058 appeared as a unique protein

product of approximately 31000 Da as estimated by comparative mobility with molecular weight standards. Cells containing APT2058 were harvested by centrifugation and inclusion bodies isolated as follows. Briefly, the cells were resuspended in lysis buffer (50 mM Tris, 1 mM ethylene diamine tetra-acetic acid (ETDA), 50 mM NaCl, pH 8.0) at 50 ml per litre of initial culture. The suspension was lysed by two passages through an Emulsiflex homogensier (Glen-Creston, Middlesex UK), followed by centrifugation at 15000 x g to purify inclusion bodies. Inclusion bodies were initially resuspended to approximately 1 $mg.ml^{-1}$ (as estimated from SDS-PAGE) in 20 mM Tris, 1 mM EDTA, 50mM 2-mercaptoethanol, pH8.5, and subsequently diluted to a final concentration of 8M urea by the addition of 10 M urea 20 mM Tris, 1 mM EDTA, 50mM 2-mercaptoethanol, pH8.5. This suspension was stirred at 4°C for 16 hours, and insoluble material removed by centrifugation at 15000 x g for 30 minutes. The APT2057 was refolded by 1 in 50 dilution into 20 mM ethanolamine , 1 mM EDTA, pH 11 buffer and static incubation at 4°C for 24 hours. Insoluble material was removed by centrifugation (10000 x g, 10 minutes), and soluble material buffer exchanged into Dulbecco's A PBS, pH 7.4 using an XK50 x 23 cm Sephadex G25 column. Refolded APT2058 was analysed by SDS-PAGE, Western blot and the effectiveness of the protein in a haemolytic assay (described in methods). Using this assay (at 1:400 dilution of human serum), the concentration of APT2058 required to bring about 50 % inhibition of lysis (IH_{50}) was approximately 3 nM.

Example 35: A method for the synthesis and characterization of APT2160 (SEQ ID NO: 48)

Compound APT2160 was generated by treating the parent compound APT2058 (at approximately 100 μM) with a three-fold molar excess of 10 mM tris-2-carboxyethyl phosphine (TCEP: in 50 mM Hepes, pH 4.5) overnight at room temperature. To this mixture was added a solution containing five molar equivalents of MSWP-1 (Example 2) in 100% DMSO for 2 hours at room temperature. APT2160 was characterized by observation of a mobility shift on non-reducing SDS-PAGE of approximately 2000 Da, consistent with the addition

of a single molecule of APT542 to APT2058. The compound was assayed in the haemolytic assay (at 1:400 dilution of human serum) and an IH_{50} value 0.03 nM was found

Example 36: A method for the synthesis and characterization of APT2184 (SEQ ID NO: 49)

Compound APT2184 is generated by treating the parent compound APT2057 with a three-fold molar excess of 10 mM tris-2-carboxyethyl phosphine (TCEP: in 50 mM Hepes, pH 4.5) overnight at room temperature. To this mixture is added a solution containing five molar equivalents of MSWP-1 in 100% DMSO for 2 hours at room temperature.

It is to be understood that the description, specific examples and data, while indicating exemplary embodiments, are given by way of illustration and are not intended to limit the present invention. Various changes and modifications within the present invention will become apparent to the skilled artisan from the discussion, disclosure and data contained herein, and thus are considered part of the invention.

This application claims priority to GB 9614871.3, filed July 15, 1996, there entirety of which is hereby incorporated by reference.

TABLE

SEQUENCE LISTING

(<- next to a peptide sequence in () signifies sequence runs C to N terminus)

SEQ ID NO:1:

GCACCGCAGTGCATCATCCCGAACAAATGCTAATAAA

SEQ ID NO:2:

AGCTTTTATTAGCATTTGTTCGGGATGATGCACTGCG

SEQ ID NO:3:

GCACCGCAGTGCATCATCCCGAACAAAGACGGTCCGAAAAAGAAGAAAAAGAAATCTCCGTCCAAATCTTCC
GGTTGCTAATAAA

SEQ ID NO:4:

AGCTTTTATTAGCAACCGGAAGATTTGGACGGAGATTTCTTTTTCTTCTTTTTCGGACCGTCTTTGTTCGGG
ATGATGCACTGCG

SEQ ID NO:5:

Gly-Ser-Ser-Lys-Ser-Pro-Ser-Lys-Lys-Lys-Lys-Lys-Lys-Pro-Gly-Asp-Cys-NH_2

SEQ ID NO:6:

Met Gln Cys Asn Ala Pro Glu Trp Leu Pro Phe Ala Arg Pro Thr Asn
Leu Thr Asp Glu Phe Glu Phe Pro Ile Gly Thr Tyr Leu Asn Tyr Glu
Cys Arg Pro Gly Tyr Ser Gly Arg Pro Phe Ser Ile Ile Cys Leu Lys
Asn Ser Val Trp Thr Gly Ala Lys Asp Arg Cys Arg Arg Lys Ser Cys
Arg Asn Pro Pro Asp Pro Val Asn Gly Met Val His Val Ile Lys Gly
Ile Gln Phe Gly Ser Gln Ile Lys Tyr Ser Cys Thr Lys Gly Tyr Arg
Leu Ile Gly Ser Ser Ser Ala Thr Cys Ile Ile Ser Gly Asp Thr Val
Ile Trp Asp Asn Glu Thr Pro Ile Cys Asp Arg Ile Pro Cys Gly Leu
Pro Pro Thr Ile Thr Asn Gly Asp Phe Ile Ser Thr Asn Arg Glu Asn
Phe His Tyr Gly Ser Val Val Thr Tyr Arg Cys Asn Pro Gly Ser Gly
Gly Arg Lys Val Phe Glu Leu Val Gly Glu Pro Ser Ile Tyr Cys Thr
Ser Asn Asp Asp Gln Val Gly Ile Trp Ser Gly Pro Ala Pro Gln Cys
Ile Ile Pro Asn Lys Cys